The Dissertation Committee for Betty Jane Kreakie Certifies that this is the approved version of the following dissertation:

Predictive Modeling of Migratory Waterfowl

Committee:

Timothy Keitt, Supervisor

Lawrence Gilbert

Lauren Meyers

Michael Singer

Zong-Liang Yang

Predictive Modeling of Migratory Waterfowl

by

Betty Jane Kreakie, B.S.; M.E.M.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August 2011

Dedication

This dissertation is dedicated to the memory of my best friend, Gracie. Every move across the country, late night study session, and much-needed break for a walk, she was contentedly there. Without her, my nearly 13-year academic journey would have felt lonely and a lot less joyful.

Acknowledgements

This research was supported by a grant to Timothy Keitt from the U.S. Department of Energy's Office of Science (BER) through the Southeastern Regional Center of the National Institute for Climatic Change Research at Duke University. Additionally, my work at the University of Texas was financially supported by teaching assistantship through School of Biological Sciences.

First and foremost, I would like to thank my advisor, Dr. Timothy Keitt. This work and my future work will be because he had trusted and believed in me. I would also like to thank Larry Gilbert, Lauren Meyers, Mike Singer, and Zong-Liang Yang (my dissertation committee) for their support and guidance. I would also like to express my appreciation to Ying Fan for her collaboration. I am not sure where my dissertation would be without her assistance.

I want to thank past and present members of the Keitt Lab, especially Jesse Lasky and Kate Behrman. Jesse and Kate read my manuscripts countless times and watched my presentations every time they were asked. Without Kate, our tiny office would have been boring and no one would have been excited about frozen yogurt.

I would like to thank my friends and family for all their love. I also appreciate their patience; I know they listened more about migratory birds than most normal people would prefer. I, especially, want to thank my husband and proofreader, Dwaine Kreakie. He reminds me that my subject and verb must always agree, and that not everything is so

serious. I also want to thank my sister, Kim Scott. She is the person to whom I can always call to cry and complain. In the end, she reminds me to put my nose down and work harder. Finally, I want to thank my parents. I do not think they ever knew exactly what I was doing, but they always supported me.

Predictive Modeling of Migratory Waterfowl

Betty Jane Kreakie, Ph.D.

The University of Texas at Austin, 2011

Supervisor: Timothy Keitt

Several factors have contributed to impeding the progress of migratory waterfowl spatial modeling, such as (1) waterfowl's reliance on wetlands, (2) lack of understanding about shifts in distributions through time, and (3) large-scale seasonal migration. This doctoral dissertation provides an array of tools to address each of these concerns in order to better understand and conserve this group of species.

The second chapter of this dissertation addresses issues of modeling species dependent on wetlands, a dynamic and often ephemeral habitat type. Correlation models of the relationships between climatic variables and species occurrence will not capture the full habitat constraints of waterfowl. This study introduces a novel data source that explicitly models the depth to water table, which is a simulated long-term measure of the point where climate and geological/topographic water fluxes balance. The inclusion of the depth to water table data contributes significantly to the ability to predict species probability of occurrence. Furthermore, this data source provides advantages over traditional proxies for wetland habitat, because it is not a static measure of wetland location, and is not biased by sampling method.

Utilizing the long-term banding bird data again, the third chapter examines the behavior of waterfowl niche selection through time. By using the methods developed in chapter two, probability of occurrence models for the 1950s and the 1990s were developed. It was then possible to detect movements in geographic and environmental space, and how movements in these two spaces are related. This type of analysis provides insight into how different bird species might respond to environment changes and potentially improve climate change forecasts.

The final chapter presents a new method for predicting the migratory movement of waterfowl. The method incorporates not only the environmental constraints of stopover habitat, but also includes likely distance and bearing traveled from a source point. This approach uses the USGS' banding bird database; more specifically, it relies on banding locations, which have multiple recoveries within short time periods. Models made from these banding locations create a framework of migration movement, and allow for predictions to be made from locations where no banding/recovery data are available.

Table of Contents

List of Tables

List of Figures

Chapter One: An Introduction to Migratory Waterfowl

Birds are one of the most thoroughly studied groups of species in ecology; this is no doubt due the intense curiosity and wonder surrounding these species (Gill 1995; Walters 2003). In particular, people have historically had a deep desire to understand the seasonal long distance migration of birds. Ancient Egyptians were some of the earliest natural historians of birds (Bleeker 1964; Velde 1980). Birds were often the subjects of their paintings and other art, and important elements of their religion were based on the natural history of birds. For example, Egyptians frequently depicted an individual's soul as a bird. Later, Linnaeus himself proposed that some bird species wintered in the tree roots while others wintered at the bottom of lakes (Alerstam and Hedenström 1998). In 1822, a white stork (*Ciconia ciconia*) was discovered in Germany with an African spear lodged in it (Bairlein 2003). This is perhaps the earliest evidence illustrating the truly remarkable distances some birds travel seasonally across the globe. The interest in the ecology of birds and, specifically, migration has by no means waned. Moreover this interest is not exclusive to professional ecologists. Today there are numerous organizations devoted to recreational bird watching, such as the National Audubon Society. Thousands of citizen scientists volunteer their time to assist with bird banding activities. Even large companies have developed to provide eco-tourism opportunities to the bird enthusiasts.

Modern ecological studies of birds have moved beyond simple interest and desire to understand this group of species to necessity. Today bird studies mainly revolve

around addressing the conservation and management issues facing these species. A large number of studies have been conducted on attempting to understand how waterfowl will be impacted by land use change; more specifically, how they will be influenced by wetland loss (Miller and Nudds 1996, Ackerman *et al.* 2006, Bartzen *et al.* 2010). Researchers have also begun to explore the potential influence of diseases (Pantin-Jackwood and Swayne 2009) and invasive species on waterfowl populations (Park 2004). Climate change, which is expected to present novel issues and exacerbate other existing ecological hazards, has begun to dominate much of waterfowl research (Sorenson, Root, and Anderson 1998, Ronka *et al.* 2005. Canepuccia *et al.* 2007). To understand the anthropogenic effects and successfully manage these species well into the future, researcher have become dependant on tracking technologies.

The development of bird banding technologies provided us with the largest increase in our knowledge about bird migration (Bairlein 2003; Bairlein 2008). Through simply tracking banding locations and dates and then the subsequent locations and dates of band recoveries, we began to understand how far birds were traveling, how quickly they moved, and what periods of the year they were traveling the greatest distances. In 1899, a Dutch school teacher, Hans Christian Cornelius Mortensen, was first to begin banding birds in a concrete effort to understand migration (Preuss 2001; Greenwood 2009). Even though not initially used to study migration, bird banding has been conducted in North America for a long time. As early as 1803, John Audubon, the famous bird naturalist and painter, banded birds to track the nesting locations of specific

individuals (Audubon, Audubon, and Coues 1994). More formalized banding in the United States did not begin until 1902 when Paul Bartsch began banding birds (Lincoln 1928; Buckley *et al.* 1998). In 1909, his work grew into the American Banding Association, which eventually evolved into the modern-day North American Bird Banding Program. The North American Bird Banding Program is a joint effort with the US Department of the Interior and the Canadian Wildlife Service. This group has already banded over 63,000,000 birds (encompassing nearly every North American bird species) and recovers about 85,000 every year.

Obviously, there have been other technological developments to track bird migration since the innovation of bird bands (Berthold and Terrill 1991). Radio tracking is one of the most straightforward and commonly-used methods to study bird movement (Marion and Shamis 1977; Mech 1983). Birds are outfitted with an identification tag and radio transmitter. Then, typically, researchers travel around with a receiver to locate the signal from a bird. The birds' identification and location are then recorded. To overcome the difficulty of locating the bird in the field, radio receivers have been placed in earth-orbiting satellites (Alerstam 1996). Most commonly the ARGOS satellites have been used for this task. Recently, birds have been fitted with Global Positioning System (GPS) devices (Hunerbein *et al.* 2000). These devices are capable of recording fine-scale movement and offer large numbers of existing software to analyze the data.

Even though these advanced technologies provide us with large amounts of much-needed fine-scale data, there are some critical limitations associated with them. These

limitations are especially restrictive when attempting to conduct large scale research over a long time period, such as the work presented here in this dissertation. Radio tracking of birds is time intensive and limited in the area that can be reasonably covered. Satellite tracking allows for a large area, yet is still costly with a limited time period of data available. GPS tracking of birds is perhaps the best available data on bird movement. Yet it is extremely expensive, and therefore the study size is greatly limited. It is also the newest method, and therefore the data only covers a short period of time and for a few individuals from a population. It is also important to note that even though there is no conclusive answer, there are concerns about the effect of carrying transmitters on the health of the study species (Gilmer *et al.* 1974; Paquette *et al.* 1997; Esler, Mulcahy, and Jarvis 2000; Wells *et al.* 2003).

Bird banding data is by far one of the simplest approaches to tracking bird movement. Even though it has its own limitations, it is still the best data available to model large scale seasonal bird migration. By relying on the USGS Banding Bird Laboratory (BBL), there is over a hundred years of data on a wide variety of species. Banding and recovery activity have a wide spatial distribution (Figure 1). This data is also collected as part of an international governmental initiative, and not by an independent researcher. This makes it more stable in the long term, and therefore increases the possible repeatability in the future. Most bird tracking data is collected for a few years, at most a couple of decades, and only provides information about one point in time. This is yet another hurdle that the BBL data overcomes.

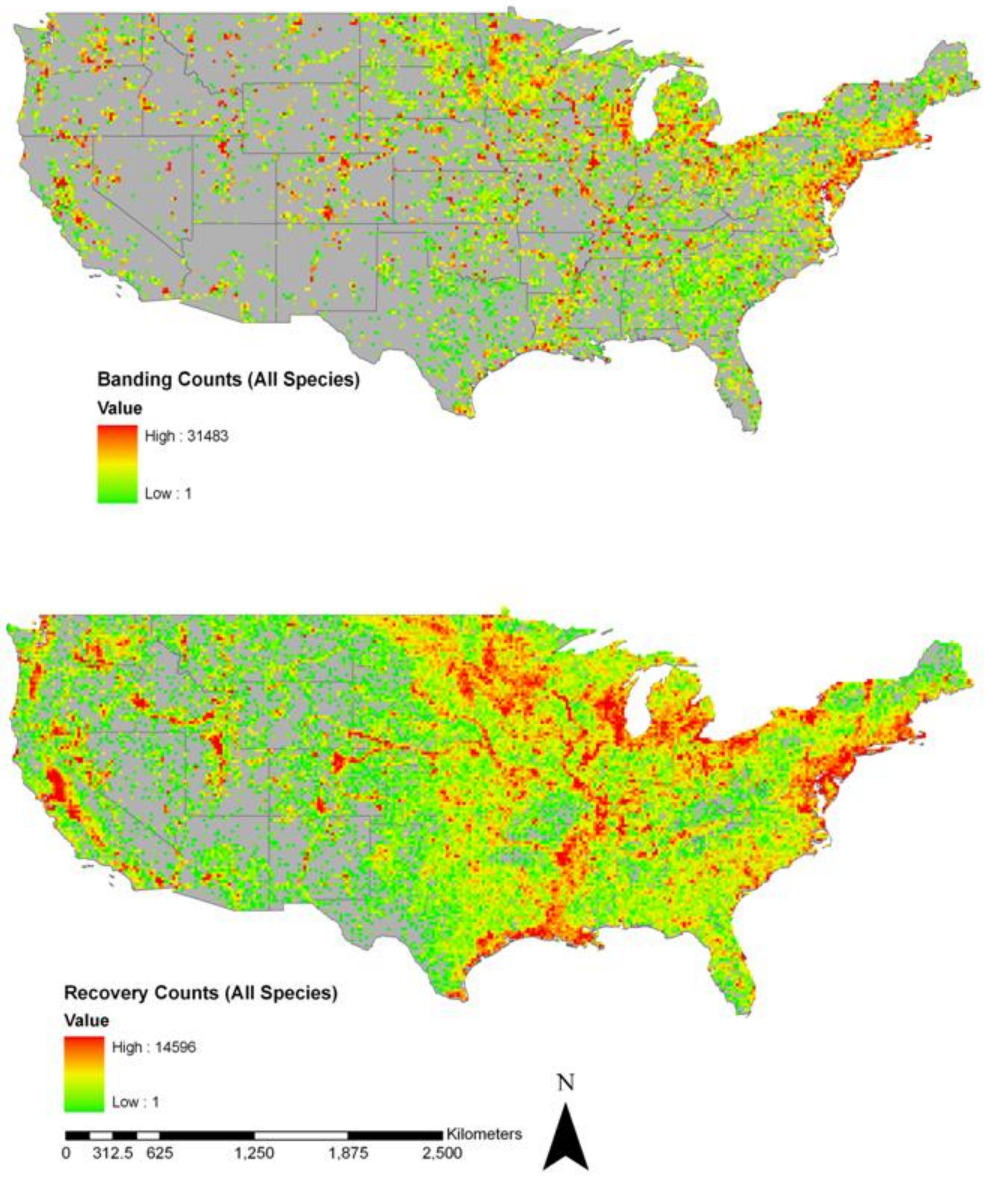

Figure 1: Maps of the banding and recovery intensity for all BBL game species in the United States

Along the migration path, it is necessary for individuals to refuel and rest before continuing on to its breeding or wintering grounds. Since birds spend a small fraction of their life in these stopover habitats and due to the difficulty of studying this portion of the bird life cycle, there is very little research conducted on stopover habitats and migration in general (Hutto 1998; Leu and Thompson 2002). This has led to a great gap in knowledge about these species, and it compromises our ability to anticipate the anthropogenic impacts on these species. The quality of wetland stopover habitat is critical in that it must provide adequate food supplies and protection from predators (Farmer and Parent 1997). Also, the amount of time spent in stopovers is much greater than initially believed (Hedenström and Alerstam 1997). As the distance between stopovers increase, birds are likely to select less suitable habitat. Therefore, as wetland habitat is lost and distance between stopovers increases, the net impact on populations is negative. The young of a species are disproportionately affected by increased distance between stopover habitat locations (Yong *et al.* 1998).

General wetland habitat quality greatly impacts waterfowl populations, regardless if it is migratory habitat or breeding/non-breeding habitat. It has been demonstrated that wetland water levels and quality is linked to waterfowl fecundity and long-term stability (Krapu, Klett, and Jorde 1983; Grier and Johnson 1988). Measures of drought intensity are also tightly correlated to waterfowl population annual counts (Sorenson *et al.* 1998). As droughts become more intense and wetland levels decrease, so do the population

numbers of waterfowl. Understanding how wetlands will respond to land-use change and climate change, will be imperative to predicting how waterfowl will respond.

In order to advance the ability to accurately predict the spatial distributions of migratory waterfowl, I have identified three key obstacles that need to be addressed. The first main hurdle that needs to be addressed is how to model species that are reliant on wetland habitat. Wetland habitat can be difficult to include in distribution modeling, since it is a dynamic habitat type and it is difficult to accurately measure. The second issue is the long-term stability of habitat selection behavior: both for breeding/non-breeding and stopover habitat. Often distribution models assume that the manner in which a species selects its habitat is static in both environment and geographic space. This assumption, often necessary due to short-term data sets, is frequently not true for migratory species. The final obstacle is how to specifically predict migratory movement of these species. As waterfowl move between their breeding and non-breeding habitat, it is difficult to predict the locations of stopover habitat, because of the complex interactions between navigational and environmental constraints. Each of these obstacles has been address in this dissertation, and is presented in turn in the following three chapters.

Chapter two of this dissertation provides a means to effectively incorporate wetland habitat, and significantly improve our ability to model the distribution of these species. This portion of my research relies on a novel data set, depth to water table (DWT), which measures the stable long-term location of the water table either at or

below this surface (Fan and Miguez-Macho 2010). The DWT takes into account the climate variables, as well as the geological and topographic variables. Since the DWT is model derived, there is the possibility to mechanistically determine the changes in the availability of wetland habitat (Franklin 2010). Additionally in this chapter, the annual cycle for each of the six study species is delineated based on species specific biology. This approach to distribution modeling will deal with species-specific seasonal requirements.

The third chapter explores the long-term changes in habitat selection of the six study species. The manner in which the study species selected habitat in the 1950s was compared to the 1990s. By doing this analysis, it was possible to gain insight about several aspects of these species' biology. For example, it is possible to examine if these species are shifting in geographic space. If there is movement in the spatial location of the species, it is possible to determine if it coincides with a shift in environmental space. If the species is not shifting in geographic location but is in environment, this may mean it is acclimating to new environmental conditions. Additionally, in this chapter I began to explore some of the potential causes for changes in environmental and geographical space. It is hypothesized that some of the underlying causes may be due to changes in phenology, population dynamics, or land-use change.

The final main chapter presents a method for predicting migratory movement. The presented technique allows for movement predictions to be made from any given point on the landscape where waterfowl habitat is likely to be located. These predictions

can be made regardless of sample intensity for the given starting location. This method makes it possible to predict where likely stopover habitat will be located from a starting location. The method incorporates not only the environmental constraints on stopover habitat, but it also incorporates the bearing and distance a species is likely to travel. This is conceivably the first work to include the interactions between environment and distance in selection of stopover habitat.

Upon beginning this project, one of the main goals was to create a collection of tools to improve our ability to conserve and manage these species. I feel that this goal was successfully achieved. By introducing a mechanistic means to include wetland habitat, it will be possible to predict how species will respond to changes in wetland. It will be possible to determine if more individuals are forced through a particular wetland due to changes in wetlands of other areas. By examining the long-term spatial stability of species distributions, it is possible to make more informed predictions about how the study species may respond. Now that there is some understanding about which species are more likely to acclimate to new environmental conditions and which are likely to track their species environmental needs, it will be possible to more accurately predict how specific species will respond to future environmental scenarios. Even though this research talks mainly about species in isolation, this work will also improve predictions made about communities of species. We can begin to examine which areas will likely be important to multiple species, and which areas when placed in conservational protection will preserve more species. The

work presented herein this dissertation will not only increase our ecological knowledge of migratory waterfowl, but will also empower us to be better stewards.

Chapter Two: Enhanced Migratory Waterfowl Distribution Modeling by Inclusion of Depth to Water Table Data[1]

ABSTRACT

In addition to being used as a tool for ecological understanding, management and conservation of migratory waterfowl rely heavily on distribution models; yet these models have poor accuracy when compared to models of other bird groups. The goal of this study is to offer methods to enhance our ability to accurately model the spatial distributions of six migratory waterfowl species. This goal is accomplished by creating models based on species-specific annual cycles and introducing a depth to water table (DWT) data set. The DWT data set, a wetland proxy, is a simulated long-term measure of the point either at or below the surface where climate and geological/topographic water fluxes balance. For species occurrences, the USGS' banding bird data for six relatively common species was used. Distribution models are constructed using diverse methods: Random Forest, GLM, and MaxEnt. Random Forest classification of habitat and non-habitat provided a measure of DWT variable importance, which indicated that DWT is as important, and often more important, to model accuracy as temperature, precipitation, elevation, and an alternative wetland measure. GLM and MaxEnt models that included DWT in addition to traditional predictor variables had significant increase in classification accuracy. Also, MaxEnt models created with DWT often had higher accuracy when compared with models created with an alternative measure of wetland

[1] A version of the material presented in this chapter is currently in review at *Plos One*

habitat. By comparing maps of predicted probability of occurrence and response curves, it is possible to explore how different species respond to water table depth and how a species responds in different seasons. The results of this analysis also illustrate that, as expected, all waterfowl species are tightly affiliated with shallow water table habitat. However, this study illustrates that the intensity of affiliation is not constant between seasons for a species, nor is it consistent between species.

INTRODUCTION

Species distribution models, especially for migratory waterfowl, are employed as a tool in diverse areas of investigation and application (Guisan and Thuiller 2005). For example, distribution models have been used to help explore how the interactions between migratory waterfowl and landscape factors will impact the spread of diseases (Gilbert *et al.* 2006; Ward *et al.* 2009; Takekawa *et al.* 2010). These approaches are used to understand how migratory birds might influence the health of the ecosystem through which they move (Post *et al.* 1998; Green, Figuerola, and Sanchez 2002). The economics of waterfowl hunting draw on distribution modeling to optimize long-term success of these game species and thus the sport (Tamisier *et al.* 2003; Casas *et al.* 2009). Perhaps the field that most heavily relies on distribution modeling is those that attempt to forecast how waterfowl will respond to anthropogenic disturbances, such as climate change (Crick 2004; Lemoine, Schaefer, and Böhning-Gaese 2007).

Given the dependency of waterfowl research and management on distribution modeling, it is critical that these tools be of the highest quality. Yet it has been shown that distribution models for birds that are migratory and have high wetland affinity are less accurate than those models for species that do not have these specific ecological traits (McPherson and Jetz 2007). Seasonal changes, in not only spatial location, but also habitat selection, contribute to this decrease in model capacity (Neave *et al.* 1996; Desgranges *et al.* 2006). Compounding temporal factors affecting model accuracy,

waterfowl are reliant on wetland habitats, which are a poorly recorded habitat type (Skagen, Granfors, and Melcher 2008). Even though there may be well-mapped modern wetland data available, due to the dynamic and often ephemeral nature of wetlands, this data will most likely be insufficient for time series analysis (Koneff and Royle 2004; Gómez-Rodríguez *et al.* 2008).

The goal of this study is to offer new strategies that will enhance distribution modeling of migratory waterfowl throughout their entire annual cycle. For each species included in our study, distribution models were constructed for each portion of the annual cycle (i.e. fall, winter/non-breeding, spring, and summer/breeding). The delineation of these events is species specific, which allows for reciprocal species specific variation in predictor variables. Even though species specific distributions were created, I utilized the availability of the community data to more accurately generate pseudo-absences when necessary (Van Der Wal *et al.* 2009). Additionally, this study introduces a novel data set to use as a predictor variable in distribution modeling of wetland species. This wetland data set, depth to water table (DWT), is a simulated long-term measure of the point either at or below the surface where climate and geological/topographic water fluxes balance (Fan and Miguez-Macho 2010).

The inclusion of wetland proxy data is a common technique used to attempt to overcome the difficulties of modeling species with high wetland affinity (Guisan and Zimmermann 2000; Guisan and Thuiller 2005). These measures range from fine scale research with direct measure of wetland quality (Milsom *et al.* 2000), through large scale

research that incorporates watershed-level hydrological modeling (Johnson *et al.* 2005; Johnson *et al.* 2010), to potentially global-scale relatively fine grain classified satellite imagery (Coops, Wulder, and Iwanicka 2009). When habitat variables have been included in distribution modeling, they are based on current classifications and not model-based prediction. In addition to the advances already made to account for wetland influence of species distribution, the inclusion of DWT data in distribution modeling provides numerous advantages to this field. The DWT data are process-driven, and allow us to investigate the underlying hydrologic drivers that may influence habitat selection. Furthermore, the DWT has a large spatial extent (nearly global) and fine resolution (approximately 270 m). The inclusion of this process-driven wetland proxy data will potentially allow us to overcome the shortcomings of forecasting future spatial distributions of countless wetland species with other approaches (like climate envelopes) (Thuiller, Lavorel, and Araújo 2005; Pearson *et al.* 2007).

The presented research is intended to augment the approaches used to construct distribution models for migratory waterfowl. Distribution models for species-specific annual cycles were assembled. This allows us to assess spatial distributions throughout the entire annual cycle, not just focusing on one portion, while adjusting for differences in timing between species. Furthermore, a novel data set, DWT, is introduced and shown to be a significant predictor variable of migratory waterfowl habitat. These data are calculations of hydrological balances between climate and geology, which will allow for more mechanistic approaches to constructing distribution models for wetland species.

Overall, the strategies presented in this research will enhance and improve distribution modeling of migratory waterfowl, and in turn allow for better management and conservation of these species.

METHODS

Species Data

The United State Geological Survey's (USGS') Banding Bird Laboratory (BBL) game bird dataset was used as the source of species presence data. Established in 1902, the BBL is a long-term monitoring project with over three million waterfowl encounter records (Buckley *et al.* 1998). The data record the incidence of banding and band-recovery events within 10-minute bins of latitude and longitude (approximately 16km). Most bands are reported by hunters and are thus terminal encounters. Each banding location and all subsequent encounter locations were treated as a known presence for that species. Owing to uncertainty in exact encounter locations, the BBL data only provide locations at 10-minute resolution; however this resolution is sufficiently fine given the broad spatial extent of the analysis. The analysis was restricted to banding and encounter events from January 1, 1990 through December 31, 1999, which was considered to be enough time to capture the main trend of defining the spatial distribution.

Six species were included in the study: American black duck (*Anas rubripes*), blue-winged teal (*Anas discors*), Canada goose (*Branta canadensis*), mallard (*Anas platyrhynchos*), northern pintail (*Anas acuta*), and wood duck (*Aix sponsa*) (Table 1). These species were chosen because they have the highest numbers of encounters and, also, had adequate sampling in all portions of the annual cycle. All six species are in the

family *Anatidae*, which are deemed typical waterfowl (Livezey 1996; Livezey 1997). Canada goose is in the subfamily *Anserinae* (geese and swan), while all other species are in the subfamily *Anatinae* (dabbling ducks).

Table 1: Season specific MaxEnt AUC scores for each study species.
The "base" variables are temperature, precipitation, and elevation. Models were constructed using the two different measures of wetland: average water table depth (DWT) from dynamically-driven hydrology model and percent wetland (PW) based on land cover classification. Asterisks indicate significantly higher AUC score when compared to the base model with other measure of wetland, and it codes the significance level: '***' 0.001 and '**' 0.01. Included is the total number of occurrence records (n) for each season of the study species.

Fall	**n**	**Base+DWT**	**Base+PW**	**DWT only**
American Black Duck	762	0.9202***	0.914	0.7274
Blue-Winged Teal	1403	0.8203	0.821	0.7444
Canada Goose	3223	0.8193***	0.8172	0.6658
Mallard	5959	0.7332***	0.7304	0.6253
Northern Pintail	314	0.8585	0.8668***	0.5881
Wood Duck	1307	0.8529***	0.8494	0.7844
Winter				
American Black Duck	765	0.9304***	0.9254	0.7649
Blue-Winged Teal	180	0.9544	0.9527	0.9153
Canada Goose	4149	0.7373	0.7354	0.6217
Mallard	4198	0.7886***	0.7814**	0.6712
Northern Pintail	1458	0.8064	0.8036	0.7368
Wood Duck	1633	0.9251***	0.9229	0.8249
Spring				
American Black Duck	18	0.9544	0.9542	0.7464
Blue-Winged Teal	18	0.844	0.8659***	0.6666
Canada Goose	225	0.7684***	0.7525	0.634
Mallard	151	0.7902	0.7856	0.5968
Northern Pintail	27	0.8331	0.8659***	0.646
Wood Duck	126	0.8784	0.8825***	0.7584
Summer				
American Black Duck	27	0.9208	0.9244	0.6942
Blue-Winged Teal	44	0.8294	0.826	0.5696
Canada Goose	2255	0.7493***	0.7433	0.6671
Mallard	651	0.8045***	0.8034	0.5888
Northern Pintail	33	0.866	0.8799	0.6048
Wood Duck	2606	0.8447***	0.8401	0.7009

Investigation of seasonal differences in distribution and habitat use required delineating the four major components of the annual cycle. BBL data was used to determine which times of the year individuals were traveling the greatest average daily distance, and these peaks in velocity were labeled as fall and spring migration. Because of the potential for confounding inter-season movement, the analysis was restricted to the mean daily traveled distance of those records where bands were recovered within 30 days of being banded. For each within-30-days recovery, the total great circle distance traveled, calculated using the "geosphere" package (Hijmans, Williams, and Vennes 2011) in R (R Development Core Team 2011), was divided by the total number of days between banding and recovery. Fall and spring migration were delineated by locating peaks in the average weekly distance traveled, and summer and winter were dated according to the appropriate intermediate seasons. These results were compared to the initial dates established by natural histories (Bellrose 1976; Bent 1987; Kaufman 1996) and the dates were adjusted when clear migration signals were present (Figure 2).

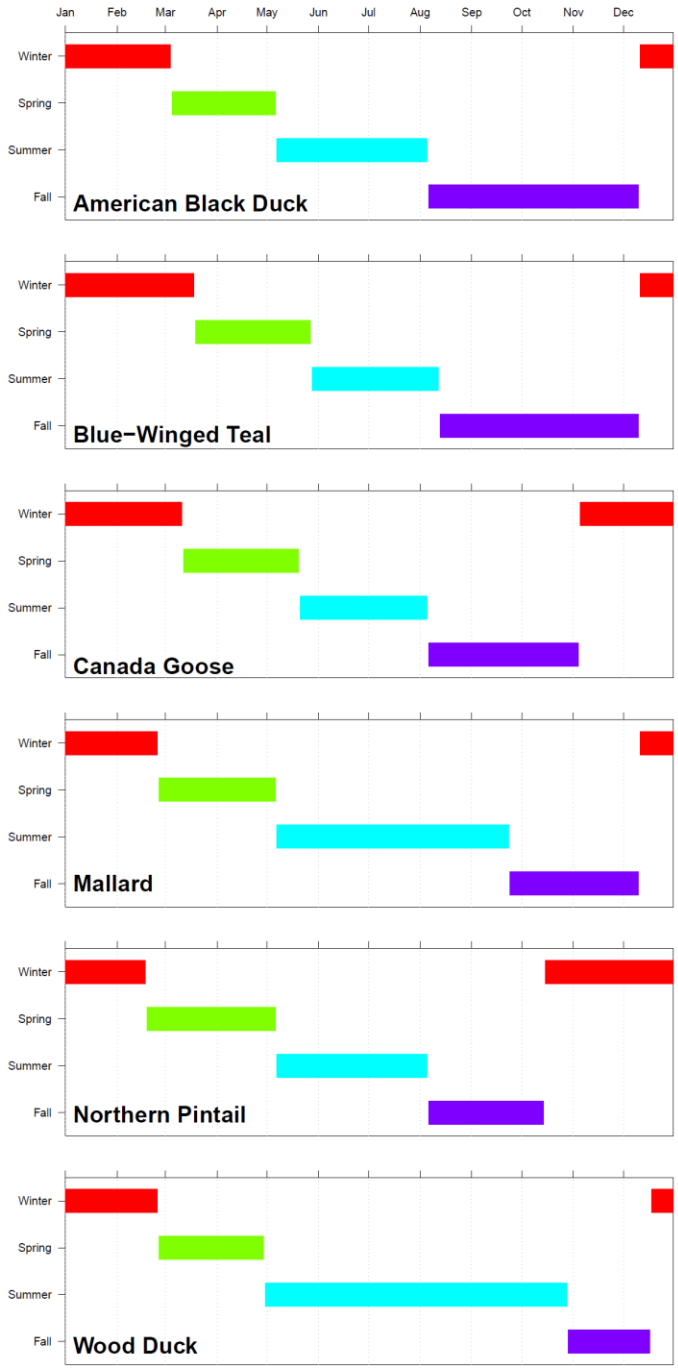

Figure 2: Barplot of annual cycle timing for study species.

Environmental Data

All predictor variables were resampled, throughout the contiguous United States study extent, to agree with the BBL data grid. Average seasonal temperature, average seasonal precipitation, and elevation were used in all distribution models. Average monthly precipitation and monthly average temperature data were obtained from the Prism Climate Group (PRISM Climate Group 2004), originally a 2.5-minute (4 km) resolution. The 3-second (90 m) Shuttle Radar Topography Mission (SRTM) Elevation Data Set was used. In addition to the three previously mentioned variables, one of two different wetland measures were included. A derived variable of percent classified wetland was created from the 2001 National Land Cover Database (NLCD) 30-meter data (Homer *et al.* 2004). The 2001 NLCD data is a land use-land cover classification of satellite, Landstat, imagery. Models were built with the inclusion of the percent NLCD wetland as a point of comparison for the models build with the DWT data.

The DWT data layer is a simulated data set that reliably predicts the location of natural wetlands (Figure 3) (Fan *et al.* 2007; Fan and Miguez-Macho 2010). The depth to water table is determined by finding the long-term stable solution of the balance between the climate-driven fluxes (precipitation and evapotranspiration) and geologic/topographic water fluxes (riverine and groundwater movement) balance. Initially, the water table was set at the surface and at each time step the modeled DWT was recalculated based on water inputs or outputs. The model was allowed to run until the water table for each cell (9-second resolution) was stable (less than 1mm change). The DWT model was validated

using 500,000+ USGS field observations of water table depth from 1927-2005, with a resulting residual mean (simulated DWT – observed DWT) of +0.443m. Fan and Miguez-Macho (2010) further tested the ability of the data to locate wetlands on the landscape. They found a strong correlation (0.8469) between field-mapped wetlands and the simulated data thresholded to 1.0m water table depth.

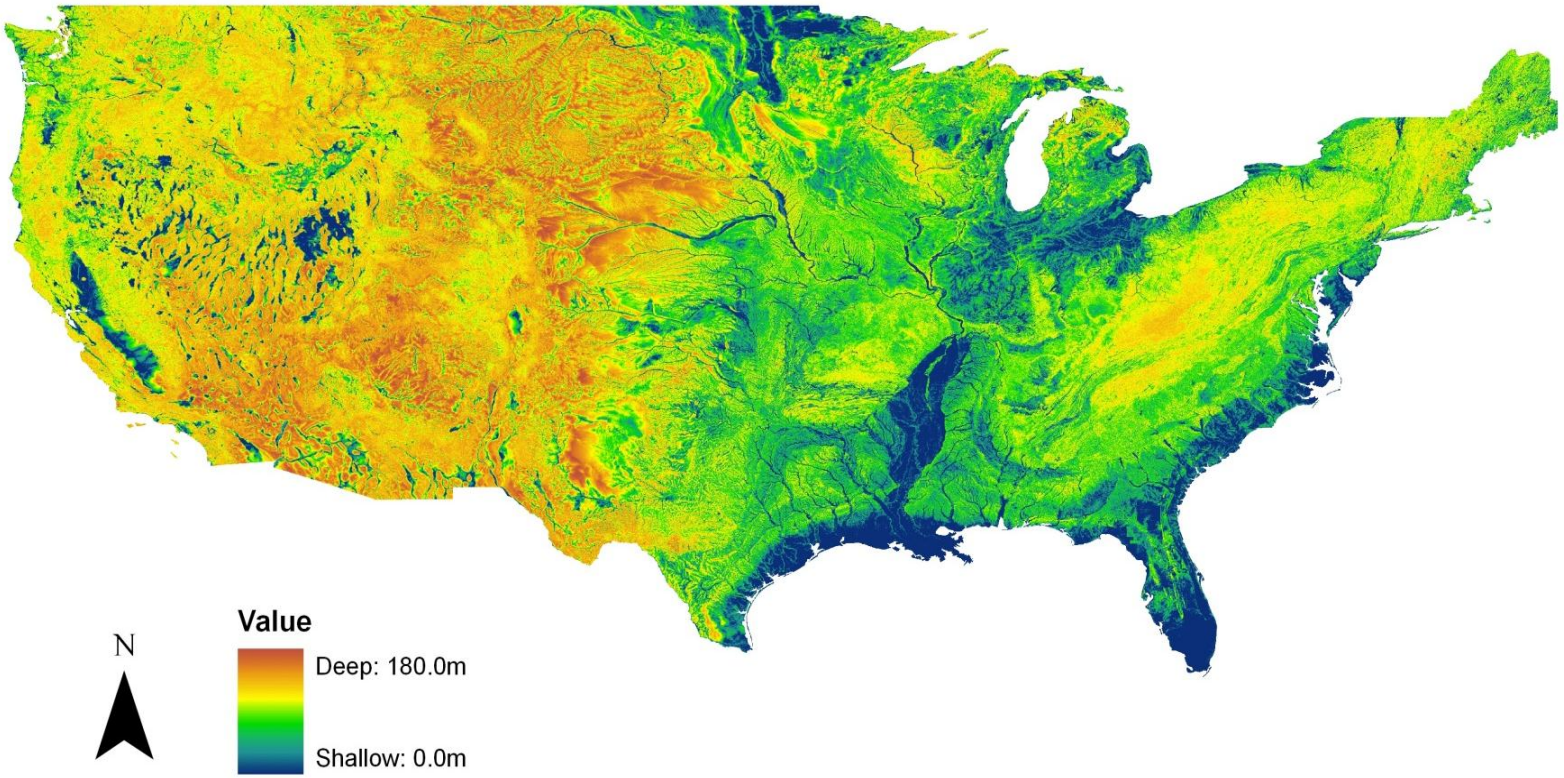

Figure 3: Map of the simulated equilibrium depth to water table for the contiguous United States. The values illustrate the depth in meters below the surface where the simulated water table is located (Fan & Miguez-Macho 2010).

Distribution Modeling and Statistical Analysis

Three methods were used to construct the distribution models for this study. Multiple methods were selected to provide results from an array of approaches each with diverse strengths and weaknesses. Random Forest was used to robustly gauge variable importance (Cutler *et al.* 2007). Random Forest is a technique that fits multiple classification trees (here, specifically 1000 trees), and outputs the mode of these individual trees (Breiman 1984; Breiman 2001). Individual trees are built with approximately 63% bootstrap of the data, and the remaining data (out-of-bag (OOB) data) used to measure its classification accuracy. Furthermore, this OOB data is used to measure variable importance. I used the mean decrease in accuracy, which is the normalized difference between classification accuracy and the accuracy when the variable values have been randomly permuted. Higher mean decrease in accuracy indicates that a variable is more important to the accuracy of the classification. All Random Forest analyses used the 'randomForest' package in R (Liaw and Wiener 2002).

Generalized linear models (GLMs) were fit using a binomial distribution and logistic link (Guisan and Zimmermann 2000; Guisan, Edwards, and Hastie 2002). Post processing stepwise regression was done to determine which variables contributed significantly to the overall model. Since the BBL data is a presence-only data set, pseudo-absences were created for both the Random Forest and GLM models. Pseudo-absences for each species and season were identified as known locations of other study species where the focal species was absent.

MaxEnt was the third and final method of distribution modeling, which is a maximum entropy approach specifically for presences-only data (Dudík, Phillips, and Schapire 2007; Phillips and Dudík 2008). It was implemented in MaxEnt 3.3.2 software package, and model set according to Phillips and Dudik (2008). Models for each season for all species were run a total of 100 times, randomizing the 70-30 training-test split of the data and the location of the background points.

Model performance of MaxEnt and GLM was measured using Area-Under-the-Curve (AUC) scores. AUC is the measure of the area under a receiver operating characteristic (ROC) curve; specifically plotting the rate of true positive classification to false positive (Manel, Williams, and Ormerod 2001; Brotons *et al.* 2004). AUC typically ranges from 0.5 (essentially random) to 1.0 (perfect fit), with 0.7 considered to be a sufficient predictor of real-world situations (Swets 1988).

The predicted probability of occurrence maps and model response curves from the MaxEnt models are presented. MaxEnt models are presented due to the fact that this method was created specifically for presence-only data, and its ability to better address the sampling bias of the BBL data (Elith *et al.* 2011). Also, only models for the winter portion of the annual cycle are presented. Winter was selected due to the relatively high sample intensity and resulting model accuracy. All other seasons are available in Appendix A.

RESULTS

Random Forest was used to measure variable importance on the accuracy of classification of presences and absences. More specifically, it was used to determine how important DWT was to the overall model and how it compared to the other predictor variables (Figure 4). Depth to water table was consistently as important as the other customary predictor variables: temperature, precipitation, and elevation. Also, DWT's importance was comparable to the importance of NLCD percent wetland for Random Forest models. The importance of DWT varied by species; with it being least important for the classification of blue-winged teal in the winter (mean decrease in accuracy = 0.097). During winter, DWT was most important for northern pintail (mean decrease in accuracy = 0.75). Comparing the importance throughout the annual cycle, DWT had the highest importance values for the spring (ranging from 1.4 for wood duck to 2.28 for American black duck).

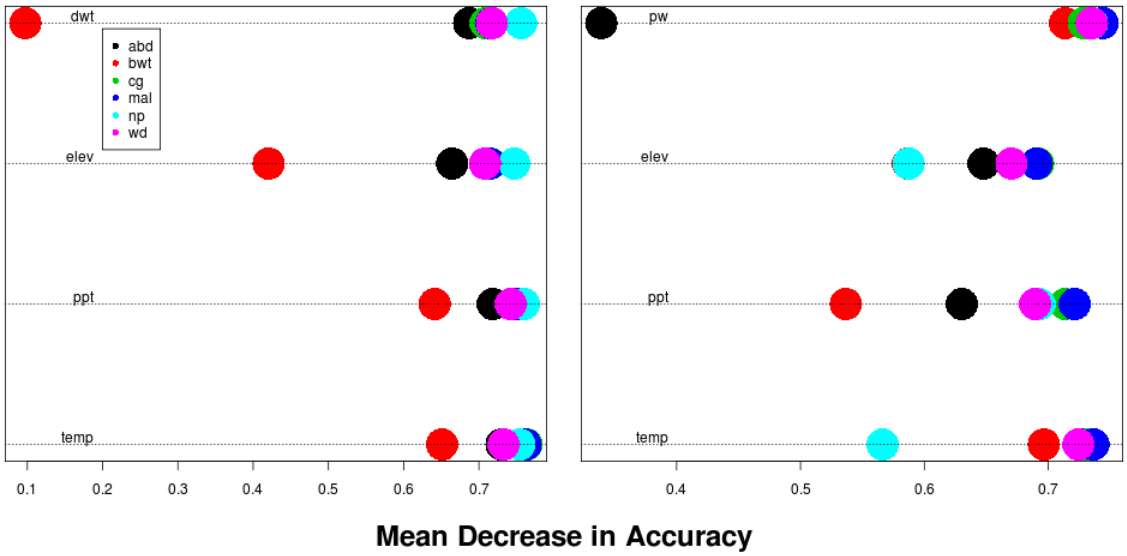

Figure 4: Plot of variable importance of winter's predicted probability of occurrence measured from Random Forest.
Variable importance is measured in mean decrease in accuracy, which is the decrease in accuracy of a classification after the variable has been randomly permuted. A higher mean decrease in accuracy means the variable contributes more to the accuracy of the classification. The abbreviations are as follows: abd (American black duck), bwt (blue-winged teal), cg (Canada goose), mal (mallard), np (northern pintail), wd (wood duck), temp (temperature), ppt (precipitation), elev (elevation), dwt (depth to water table), and pw (NLCD's percent wetland).

All MaxEnt and GLM models created with a wetland variable (DWT or NLCD percent wetland) had significantly higher AUC score than those models (from here forward referred to as "base models") created with only temperature, precipitation, and elevation. When the AUC of MaxEnt models for each species were directly compared between base with DWT and base with percent wetland, 11 MaxEnt models build with DWT had a higher AUC than percent wetland model for that species (Table 1). Six

MaxEnt models had no statistical difference between AUC scores, and 4 MaxEnt base and percent wetland models had higher AUC scores than the reciprocal base and DWT model. Nine of the 24 MaxEnt models built with only DWT as predictor variable had an AUC over 0.7.

For GLM models with stepwise variable selection regression, DWT was selected as a significantly contributing variable ($p < 0.01$) for 16 of 24 models (Table 2). Mean percent wetland was selected, at the same level of significance, for 12 of 24 models. When the GLM models for each species were compared between base with DWT and percent wetland, 7 out of 24 models had a higher AUC with DWT than those built with percent wetland. Four of 24 GLM models built with only DWT had an AUC over 0.7.

Table 2: Season specific GLM AUC scores for each study species.
The "base" variables are temperature, precipitation, and elevation. Models were constructed using the two different measures of wetland: average water table depth (DWT) from dynamically-driven hydrology model and percent wetland (PW) based on land cover classification. Asterisks indicate the significance of the variable in the regression model. Significance codes: '***'< 0.001, '**' <0.01, and '*' <0.05.

Fall	**Base+DWT**	**Base+PW**	**DWT only**
American Black Duck	0.8129**	0.8281***	0.6051
Blue-Winged Teal	0.6978***	0.6895***	0.6679
Canada Goose	0.7068***	0.7382***	0.4952
Mallard	0.6838***	0.904	0.6134
Northern Pintail	0.7641***	0.9336	0.548
Wood Duck	0.8312	0.9042**	0.7124
Winter			
American Black Duck	0.7935*	0.7945*	0.6315
Blue-Winged Teal	0.9085***	0.9050***	0.8645
Canada Goose	0.729	0.7330***	0.6236
Mallard	0.6386**	0.6439***	0.5272
Northern Pintail	0.6275***	0.6200***	0.6324
American Black Duck	0.8721***	0.8795***	0.7447
Blue-Winged Teal			
Canada Goose			
Mallard	0.8685**	0.8803	0.4246
Northern Pintail	0.7760**	0.7425	0.6516
Wood Duck	0.9408**	0.9404*	0.9342
Mallard	0.8451	0.8518*	0.8227
Northern Pintail	0.7816	0.7986*	0.6875
Wood Duck	0.8877**	0.8863	0.6605
Summer			
American Black Duck	0.7117*	0.7623***	0.6141
Blue-Winged Teal	0.7314**	0.7279	0.6916
Canada Goose	0.6084***	0.6085	0.5354
Mallard	0.7612	0.7629***	0.5872
Northern Pintail	0.9137	0.9130*	0.6275
Wood Duck	0.7010***	0.7145***	0.5857

The MaxEnt winter predictions, created with base predictor variables and DWT, for all species are presented in Figure 5 (all other MaxEnt predictive maps are provided in Appendix A). Predictions are in line with the fact that all species should be centered in the southern portions of their ranges. The highest predicted values for blue-winged teal, northern pintail, and wood duck are along the southern portion of the east coast and the Gulf of Mexico coast up the Mississippi River. Canada goose and mallard, to a lesser degree, have large areas of mid-range predicted probability of occurrence in the central portion of the United States. American black duck's predictions are focused in the northeastern portion of the country, while avoiding the peaks of the Appalachian Mountains. All species have some moderate predictions along the west coast, especially in the Central Valley in California.

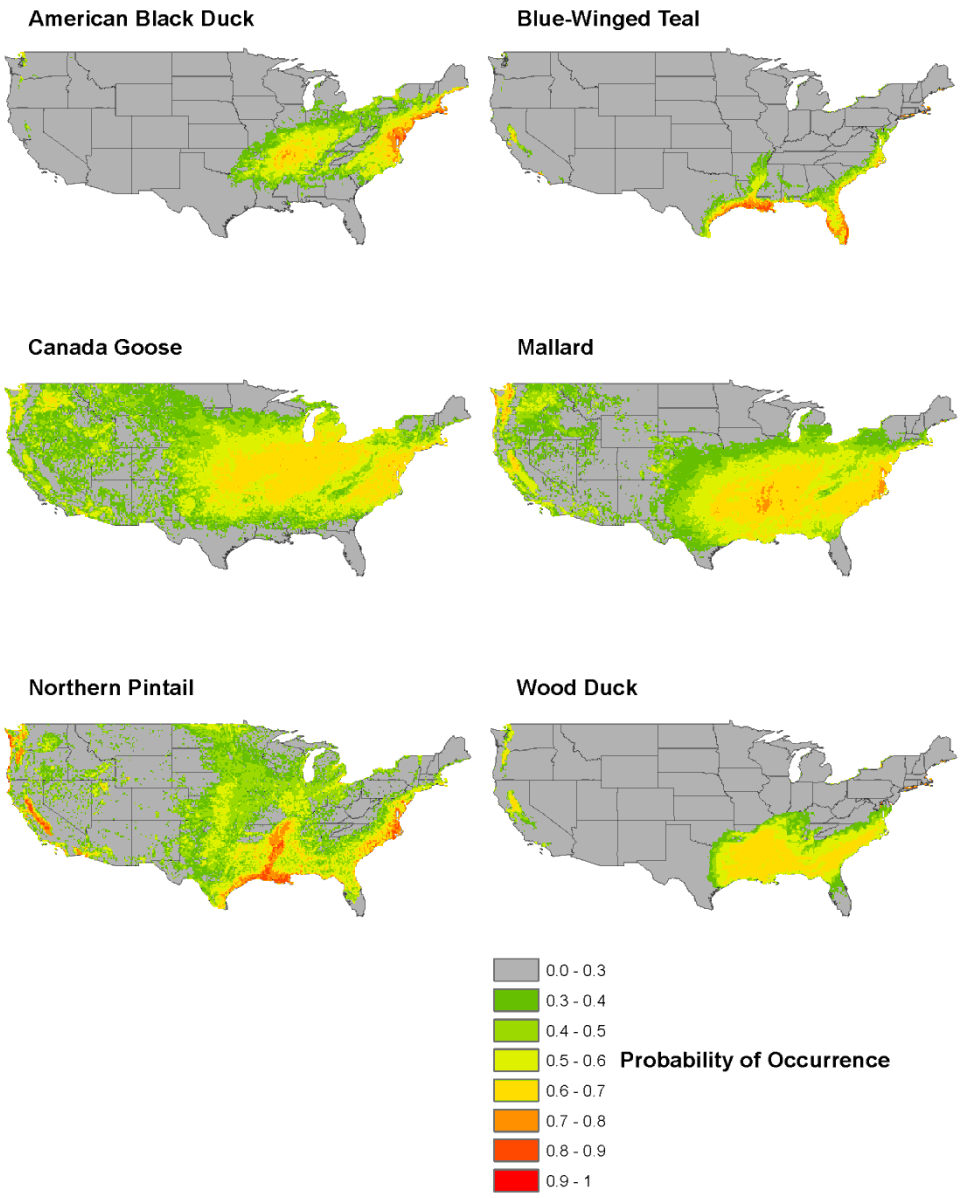

Figure 5: Maps of predicted probability of occurrence for all study species' winter habitat. Predictions were created using MaxEnt with 100% of known presence locations to increase accuracy of the visual representation. Temperature, precipitation, elevation, and water table depth were the predicted variables used to construct the probability surfaces.

For each of the study species in the winter, the relationship between DWT and MaxEnt predicted probability of occurrence is presented in Figure 6 (all other response curves are provided in Appendix A). The distribution of each species is skewed towards the shallow water tables. Canada goose and mallard's distribution are less skewed to the left than the other species; they have a more gradual decrease in predicted occurrence as the water table becomes deeper. Northern pintail has the highest peak at 0.6 at the shallowest water table level.

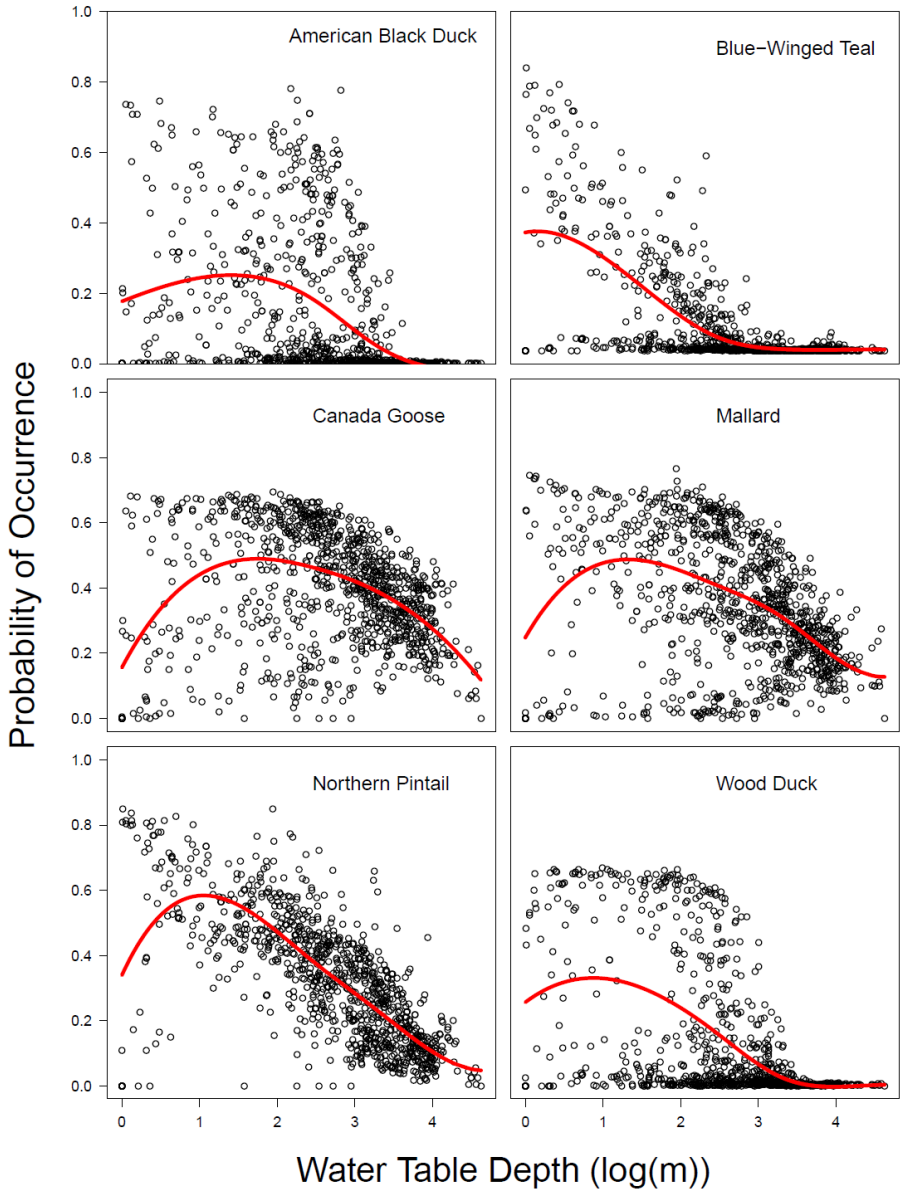

Figure 6: Plot of the relationship between the log transformed water table depth (m) and predicted probability of occurrence for each study species in winter.
The plots were constructed by selecting 1,000 random points from the predicted probability of occurrence surface. The red curve is a smoothing spline fit to the mean of the data points, and meant to illustrate the trend of the data.

DISCUSSION

The foremost goal of this study was to present strategies that would enhance our ability to create accurate distribution models of migratory and wetland species. By modeling distributions based on species-specific annual cycles and introducing a novel data set, I was able to successfully accomplish this goal. All species, in all four portions of the annual cycle, had MaxEnt models (base variables plus DWT) with AUC scores greater than the 0.7 threshold. Additionally, it was possible to show that the DWT data set significantly contributed to the distribution models of these species. This was illustrated, first, by showing that DWT was consistently ranked higher in variable importance for Random Forest classifications. Secondly, DWT significantly added to the classification accuracy of MaxEnt models when compared to models created with only temperature, precipitation, and elevation. Also, step-wise variable selection in the GLM models reliably selected DWT as a significantly contributing variable. And finally, the DWT data set performed as well, and often better than, as a standard proxy for wetland habitat, classified satellite imagery.

The DWT data offers more advantages to distribution modeling beyond the increased model performance presented here in this research, most of which are due to the fact that it is a model-derived data set. Most importantly the DWT has the potential to be more than a static measure of wetland habitat. The DWT is a measure of the point where hydrologic, topographic, geologic, and climatic fluxes balance. By predicting how the depth to water table changes according to changes in the environment, like climate

change, it will allow for more mechanistic predictions of how wetland species will respond. This data set avoids many of the biases that are present in the more traditional measures of wetland habitat quality or quantity. For example, if using mapped delineated wetlands, especially for studies at the continental scale, there is a concern that all those who did the delineation were using the same definition of a wetland (Stevens and Jensen 2007). Additionally for studies of large spatial extent, there are often large gaps in digitally available mapped wetland data. These concerns are also true for classified satellite imagery. Often it is unclear, if what is being classified as wetland is truly wetland on the ground (Gonzalez and Hollister 2008).

One concern with the DWT data set for distribution modeling is its accuracy at finer scale. The model from which the DWT data are derived does not for example incorporate detailed data on local water extraction and management. Water levels in many wetlands (and wildlife refuges in particular) are actively managed and therefore are expected to deviate from the DWT data. At the relatively coarse 10-minute scale of this study, these deviations are likely not of great concern as the hydrological model will generally identify low-lying areas where water accumulates. These are the same areas where both managed and unmanaged wetlands will predominantly occur. At finer sub-kilometer scales, the limitations of the modeling approach might however become much more apparent as even small changes in water table could be the difference between wetland habitat and dry ground. We are currently investigating the performance of the

DWT data for wetland-species distribution models using fine-scale species occurrence data (Kreakie and Keitt, unpublished data).

In addition to the methodological advantages presented here, this research provides insight into the ecology and behavior of these six species. Each species responds differently to the hydrologic regime, even within the group of waterfowl (Bolduc and Afton 2008). By using the response curves (Figure 6), it is possible to quantitatively gauge how each species will respond to the changes in the depth to water table. All six study species have increased predicted probability of occurrence toward shallow depth to water table, but these distributions are not uniform between species. For example, both American black duck and blue-winged teal are more skewed toward the shallow end of depth to water table than Canada goose and mallard. The more uniform predicted probability of occurrence across the range of DWT for mallard and Canada goose could be due to multiple factors. First, this could be due to true behavior of these species. These two are more generalist species, and can often been seen in areas devoid of wetlands (like golf courses and agricultural fields). Second, this uniform predicted response to DWT could be due to the 10-minute scale of the analysis. Canada goose and mallard prefer to be in wet habitat, but are also fine with wet areas nested within an area of relatively deep DWT (for example, a housing subdivision's retention pond) (Conover and Chasko 1985).

It is not only possible to compare between species response to DWT, but it is also possible to examine how the predicted responses to DWT changes between season

(Figure 6 and Figure A5-A7). For example, blue-winged teal is tightly constrained to shallow DWT in the winter. However, this predicted behavior changes in the summer/breeding season. I hypothesize that conceivably blue-winged teal is foregoing wetland habitat for drier, and perhaps safer, upland nesting sites. It is also important that this is may be another relic of the 10-minute scale. The breeding area of blue-winged teal is concentrated in the Prairie Pothole region, where there are numerous small wetlands within a relatively dry upland landscape matrix.

This study illustrated that the new process-driven depth to water table data set can be used as a significant predictor variable in distribution modeling of migratory waterfowl. The depth to water table data set is new and has some important hurdles to overcome, such as how to effectively handle human manipulation of the water table. Yet, despite being in its early period, the future research possibilities are abundant and exciting. To date, forecasting the response of wetland species to climate change has been severally limited due to the dynamic nature of wetlands. This issue becomes compounded when considering migratory species that rely on wetlands for stopover habitat. It becomes nearly impossible to make predictions about the future of migratory waterfowl and how manage accordingly, when there has been no mechanistic means to forecast key wetland habitat across the entire migration route. The DWT data will allow for us to begin to move beyond these obstacles, and make more vigorous prediction about the future of migratory waterfowl.

Chapter Three: North American Waterfowl Distributions: Analyzing Long-Term Stability and Case Studies Exploring Potential Drivers of Change[2]

ABSTRACT

Knowledge about species distribution variations through time increases basic ecological understanding, improves species management and conservation, and allows for enhanced predictions about the future of the species. Research regarding changes in distribution through time is difficult, especially for migratory waterfowl, due to data limitations. The main goal of this research is to determine how the distributions of six migratory waterfowl species have changed from the 1950s to the 1990s. This goal is obtained by overcoming the data limitations of long-term distribution analysis. I used the USGS' banding bird laboratory data, which provides over a hundred years of banding and recovery data. For the wetland habitat variable, a novel mechanistic measure of depth to water table was used. Distributions of all six study species, in all portions of the annual cycle, varied uniquely. These changes were unique in both geographic and climate space. Case studies were used in an attempt to tease apart some of the reasons why the distributions behaved in certain manners. In particular, the fall distributions of blue-winged teal and wood duck were examined in more detail. It was hypothesized that the distribution changes could be a result of fall migration phenology, population dynamics, or land use change. The main conclusions are that distributions of migratory waterfowl are not static over this study period, and that the manner in which the distribution change and, therefore, the underlying drivers of change are species and season specific.

[2] A version of the material presented in this chapter is currently in prep for submittal to *Diversity and Distributions*

INTRODUCTION

Understanding the relationship between species occurrences and the environment is foundational to ecology (Grinnell 1917; Andrewarthe and Birch 1954; MacArthur 1972). This relationship modeled explicitly and spatially is summarized in species distribution models (Guisan and Zimmermann 2000; Guisan and Thuiller 2005). Once we know why a species occurs where it does, it is also as important to understand how the species distribution changes over time. Known relative rates of change in climate and geographical space provide insight into the adaptability of a species to shifting environmental conditions (Hulme 2005) or the migratory ability of the species to track suitable climatic conditions (Iverson and Prasad 1998). Adaptability and migration potential will increase a species long-term resilience (Parmesan 2006; Visser 2008).

Knowledge about the long-term stability or variation of a species distribution has numerous applied uses. It allows us to understand how species have responded to past climatic changes (Acevedo *et al.* 2010; Dudei and Stigall 2010; Rubidge *et al.* 2011), and predict how they may respond to future variation in the climate (Pearson and Dawson 2003). This information makes it possible to more accurately predict how invasive species (Loo, Nally, and Lake 2007) and diseases (Parham and Michael 2010) move across the landscape. The overall general health of a species may be ascertained by examining the spatial distribution over time (Bartel and Sexton 2009; Graham *et al.* 2010). A stable or increasing distribution could be an indicator of a species that is viable

or growing, where a decrease in range might be a warning sign for a struggling population (Holt 1992; Thomas *et al.* 2001).

Even though it is critical to understand how species distributions change over time, this type of research is difficult to conduct (Guisan and Zimmermann 2000; Araújo and Pearson 2005). Typically, data limitations prevent researchers from even attempting this work for waterfowl (Hitch and Leberg 2007). There are very few species occurrence data sets that span more than a few years; whereas, depending on the species, it may require decades of data to provide enough information. Data limitations are not exclusive to occurrence data. Often key environmental variables of species distribution are unavailable over long time periods. Specifically related to this research, delineated or classified wetland data for a long time period and across a large spatial extent is not available. Without wetland habitat information, it is nearly impossible to say anything concrete about how wetland species' have responded to historical changes in their habitat (Randin *et al.* 2006). Furthermore, predictions about how they respond to future changes to the environment will be based on incomplete understanding of the species (Heikkinen *et al.* 2006).

The main goal of this research is to determine how the distributions of six migratory waterfowl species have changed from the 1950s to the 1990s. This investigation includes where the species have adjusted in spatial location, and how they have changed in climate space. In other words, even if the species have not changed the extent of their spatial distribution, I explored how the established area may have changed

in its environmental characteristics. The goal was accomplished by overcoming the data limitations previously mentioned. The most difficult hurdle, obtaining dynamic wetland habitat data, was surmounted by utilizing depth to water table (DWT) data. The DWT data is a model-derived data set that mechanistically determines the point where geological and hydrological fluxes balance (Fan et al. 2007; Fan and Miguez-Macho 2010). For species occurrence data, the USGS' banding bird laboratory (BBL) data was used, which is a long-term record of banding and subsequence band recovery locations (Buckley *et al.* 1998).

The presented research also includes three case studies, which explore the potential underlying drivers of change to the distributions. The intent of these case studies is to provide more detail about these specific examples. Moreover they are intended to show that how species respond to changed is not uniform, even between similar species. The abiotic factors impacting species distributions are numerous, and the manners in which species respond are as abundant. The case studies also illuminate that investigating the changes in species distributions over time is a hypothesis generator for future ecological research. This avenue of research provides immediate insight into species responses to climate change, and also provides direction for future avenues of research.

The aim of this study is to determine changes in geographic and/or climate space of six migratory waterfowl species. Three case studies are presented to illustrate potential drivers of change in species distributions. The presented work is an attempt to

move beyond the false assumption that species distributions are static in time, and that species respond to changes in the environment uniformly.

METHODS

Species Data

The USGS' Bird Banding Laboratory (BBL) data was the source of species occurrence data, which are data regarding the banding and subsequent recoveries of birds (Buckley *et al.* 1998). The BBL data dates back to 1902, when founder, Paul Bartsch, developed a scientific method for banding birds, and initiated the project by banding over a hundred birds himself. The American Bird Band Association began in 1909, and the banding effort has since increased exponentially. There have been approximately 3.5 million recoveries to date, and about 1.2 million birds are currently banded. With over a hundred years of banding and recovery data and a considerable sample size, the BBL database was a logical choice for a study attempting to understand the changes in species distribution over time.

Six migratory waterfowl species were selected for this study: American black duck (*Anas rubripes*), blue-winged teal (*Anas discors*), Canada goose (*Branta canadensis*), mallard (*Anas platyrhynchos*), northern pintail (*Anas acuta*), and wood duck (*Aix sponsa*). These six species were chosen due to their consistent sample sizes throughout the study periods. Presence locations (specifically, BBL recovery locations) of the study species are reported according to a 10-minute (approximately 16km) grid of latitude and longitude. The BBL 10-minute grid is sufficiently coarse enough to account for uncertainty in exact location, yet appropriate for my study due to its large spatial extent.

For the 1950s and 1990s study periods, distribution models were created for each portion of the study species' annual cycle: winter/breeding, spring migration, summer/non-breeding, and fall migration. The timing of annual cycle was species specific, and created according to the methods described in Kreakie *et al.* (*in review*).

Environmental Data

Four predictor variables were used to construct the species distribution models: temperature, precipitation, elevation, and depth to water (DWT). All predictor variables were resampled to correspond with the BBL 10-minute grid for the entire spatial extent for this study, the contiguous United States. Monthly temperature and precipitation, originally at 2.5-minute (4 km) resolution, were obtained from the Prism Climate Group (2004), and averaged over species specific season. The 3-second (90 m) Shuttle Radar Topography Mission (SRTM) Elevation Data Set was used.

The DWT data set, originally 270m resolution, is a model-derived proxy for wetland habitat. The DWT data have been shown to have high correlation with field-mapped wetlands (0.8469) (Fan and Miguez-Macho 2010), and a significant predictor variable of migratory waterfowl distributions (Kreakie, Fan, and Keitt *in review*). The presently available DWT data set is a measure of the long-term stable location of the water table. I used data 42 USGS' groundwater wells, which reported adequate sampling in both study periods, to determine the deviation of the water table from the average depth to water.

Species Distribution Modeling and Statistical Analysis

The relationship between species occurrences and the environmental variables was modeled using MaxEnt (Phillips, Anderson, and Schapire 2006; Phillips and Dudík 2008), which is a machine learning algorithm that utilizes maximum entropy theory. MaxEnt was selected over other methods due to the fact that it is a species distribution approach designed to handle presence-only data, which the BBL data are (Elith *et al.* 2011). The MaxEnt 3.3.2 software package was used, and set model parameters based on Phillips and Dudik (2008). One hundred iterations for each model were run; in order to randomize the 70-30 training-test split and the background points' locations.

The model's performance was evaluated using the area under the receiver operating characteristic, known as AUC. The AUC specifically measures the area under the curve plotting the true positive rate versus the false positive rate (Manel, Williams, and Ormerod 2001; Brotons *et al.* 2004). AUC scores range from 0.5 (essentially random) to 1.0 (perfect fit), with 0.7 considered to be a sufficient predictor of real-world situations (Swets 1988).

A kappa statistic was calculated to determine the percent agreement between the 1950s and 1990s outputs (for each species and season) (Bell and Fielding 1997). While taking chance agreement into account, the kappa statistic essentially encapsulates all the information supplied in a confusion matrix of a presence/absence classification. The threshold probability of 0.6 was used to determine presence/absence. Kappa was

calculated using the "PresenceAbsence" package (Freeman 2007) in R (R Development Core Team 2011). A kappa statistic greater than 0.75 is considered to be a good agreement between the two data sets, and below 0.4 is poor (Landis and Koch 1977).

In interest of brevity, only the fall spatial and climatic results are presented and discussed here. However, all seasons' results are provided in Appendix A.

RESULTS

Number of presences used to build the species distribution models ranged from 25 individual recoveries (1990s summer northern pintail) to 6,185 recoveries (1990s winter mallard) (Table 3). Wood duck had, by far, the largest increases in sample size from the 1950s to the 1990s, which range from 216% increase to 603%. Five of the 48 models constructed for this study had AUC scores below the 0.7 threshold; four of the five were mallard and the other Canada goose, which are more habitat generalists than the other study species (Bellrose 1976; Bent 1987; Kaufman 1996). The best performing model was American black duck 1950s spring with an AUC of 0.965.

Table 3: MaxEnt performance results for 1950s and 1990s species distribution models. Number of occurrences (n) for each species in all portions of the annual cycle is provided. The AUC score, with its corresponding standard deviation in parentheses, are also present.

Fall

	n	1950s	n	1990s
American Black Duck	1730	0.874 (0.0058)	921	0.904 (0.0068)
Blue-Winged Teal	1462	0.841 (0.0096)	1421	0.807 (0.0102)
Canada Goose	1236	0.791 (0.0114)	4816	0.794 (0.0054)
Mallard	6544	0.662 (0.0065)	6172	0.678 (0.0065)
Northern Pintail	751	0.878 (0.0127)	313	0.851 (0.0257)
Wood Duck	875	0.884 (0.0078)	2769	0.817 (0.0062)

Winter

	n	1950s	n	1990s
American Black Duck	1054	0.906 (0.0059)	871	0.916 (0.0062)
Blue-Winged Teal	156	0.929 (0.0234)	296	0.950 (0.0103)
Canada Goose	2304	0.710 (0.0111)	6057	0.695 (0.0065)
Mallard	4706	0.681 (0.0077)	6185	0.678 (0.0066)
Northern Pintail	2862	0.719 (0.0094)	1568	0.774 (0.0118)
Wood Duck	528	0.920 (0.0081)	2755	0.842 (0.0056)

Spring

	n	1950s	n	1990s
American Black Duck	152	0.965 (0.0108)	45	0.946 (0.0287)
Blue-Winged Teal	70	0.827 (0.0396)	27	0.821 (0.0661)
Canada Goose	142	0.768 (0.040)	1191	0.813 (0.011)
Mallard	455	0.777 (0.022)	671	0.802 (0.0175)
Northern Pintail	156	0.836 (0.0373)	45	0.896 (0.0461)
Wood Duck	31	0.919 (0.0504)	218	0.882 (0.0197)

Summer

	n	1950s	n	1990s
American Black Duck	71	0.958 (0.0131)	36	0.949 (0.031)
Blue-Winged Teal	56	0.872 (0.0553)	32	0.816 (0.0868)
Canada Goose	156	0.836 (0.0386)	1545	0.794 (0.0105)
Mallard	373	0.814 (0.0235)	961	0.796 (0.0133)
Northern Pintail	64	0.837 (0.0576)	25	0.837 (0.0761)
Wood Duck	412	0.919 (0.010)	2639	0.818 (0.0064)

Kappa statistics ranged from 0.0867 for Canada goose in the summer to 0.841 for American black ducks in the spring (Table 4). American black duck had the highest average kappa statistic across all season (i.e. least change between 1950s and 1990s across seasons) of 0.744. Canada goose had the lowest kappa average (0.293) across all seasons.

Table 4: Kappa statistics comparing all 1950s and 1990s distribution models. The corresponding standard deviation is in parentheses.

	Fall	Winter	Spring	Summer
American Black Duck	0.790 (0.010)	0.736 (0.014)	0.841 (0.013)	0.607 (0.019)
Blue-Winged Teal	0.399 (0.010)	0.798 (0.013)	0.614 (0.011)	0.373 (0.010)
Canada Goose	0.375 (0.010)	0.150 (0.008)	0.562 (0.009)	0.0867 (0.008)
Mallard	0.469 (0.010)	0.736 (0.008)	0.569 (0.019)	0.596 (0.010)
Northern Pintail	0.731 (0.009)	0.554 (0.010)	0.435 (0.013)	0.450 (0.012)
Wood Duck	0.598 (0.010)	0.489 (0.011)	0.100 (0.011)	0.462 (0.012)

Whereas Table 4 reports the absolute difference between predicted surfaces for the 1950s and the 1990s, Figure 7 illustrates where there were the largest increases or decreases in fall predicted probability of occurrence in the 1990s compared to the 1950s. Most strikingly, wood duck had the largest area of significantly increased predicted probability of occurrence for fall of the 1990s compared to fall of the 1950s. Even with a vast area of increase, the wood duck had patchy decreases in predicted probability (e.g. in Texas along the Gulf of Mexico coast). American black duck has the lowest total area with significant changes in predicted probability of occurrence. Mallard and Canada goose, to some degree, illustrated similar changes in location of their greatest increases

and decreases in predicted probability; patchy decreases in the central portion of the United States and increases along the New England coastline through the Hudson River Valley. Blue-winged Teal had large decrease in the north central United States, and large increase along the Gulf of Mexico coast and the southern end of the Mississippi River.

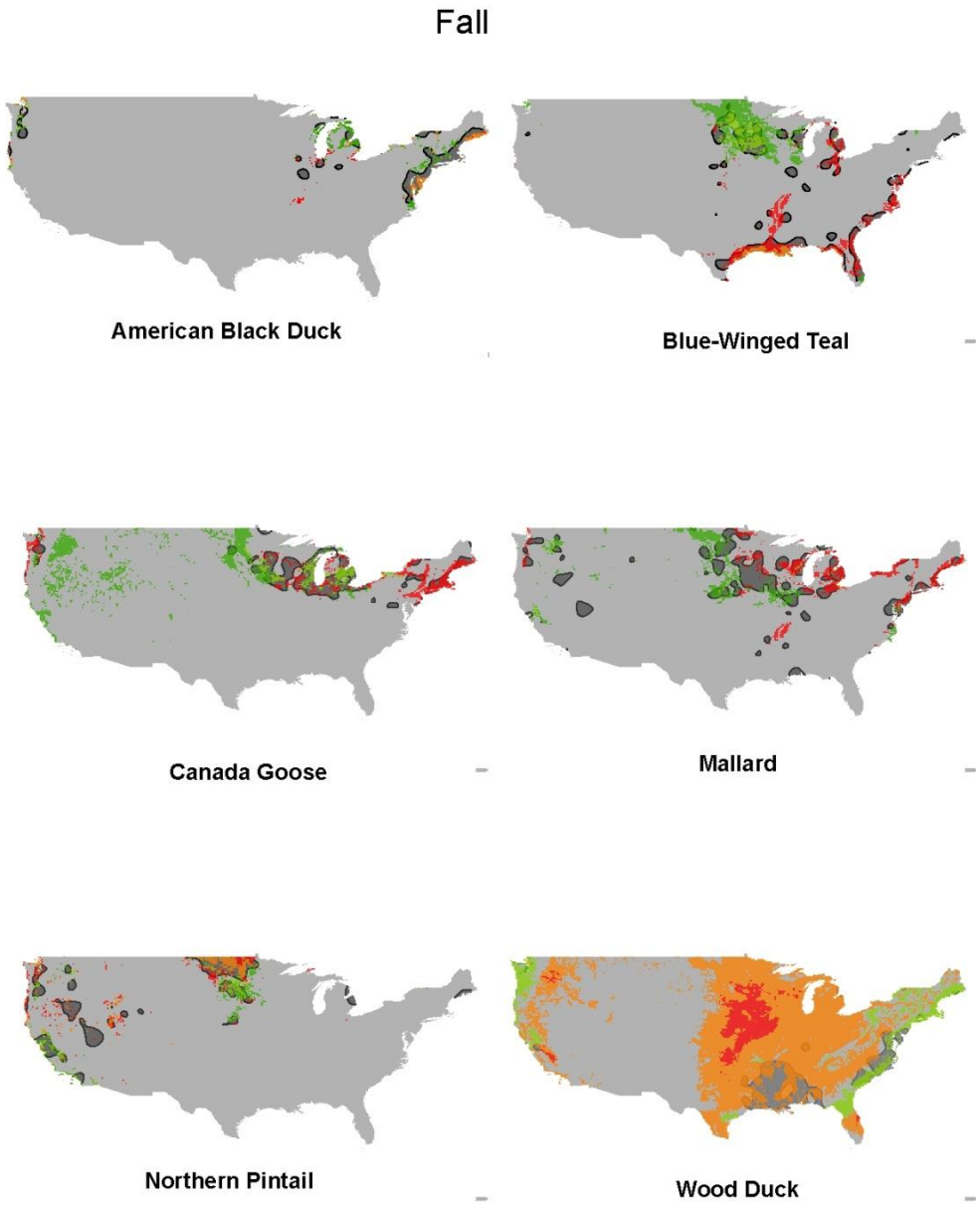

Figure 7: Maps of changes in fall 1990s predicted probability of occurrence relative to the fall 1950s predicted probability of occurrence.
The dark gray polygons encircle 1990s optimal habitat; areas with 60% or greater predicted probability of occurrence. The red illustrate areas that had a 25% or greater increase in predicted probability of occurrence in the 1990s compared to 1950s. The orange area ranges from 10-25% increase in 1990s. The dark green areas are where there was a 25% or greater decrease in predicted probability, and light green from 10-25% decrease.

CASE STUDY ONE: POTENTIAL FALL PHENOLOGICAL SHIFT OF BLUE-WINGED TEAL

Based on the spatial changes in predicted probability of occurrence (Figure 7), it was hypothesized that in the 1990s blue-winged teal may be leaving the northern breeding grounds earlier in comparison to the 1950s. The following case study is an attempt to explore this hypothesis further.

Case Study One: Methods

To determine if fall migration phenology had changed from the 1950s to the 1990s, it was necessary to determine when birds in each study period were traveling the greatest directed velocity. For this case study, I only used blue-winged teal that had been recovered within 30 days of being banded. This ensured that confounding effects of inter-seasonal movement were avoided. The great-circle distance (calculated using the "geosphere" package in R (Hijmans, Williams, and Vennes 2011) traveled for each within-30-days blue-wing teal was calculated. This distance was divided by the total number of days between banding and recovery, which provided an approximation of the bird's velocity. Finally, the average velocity by week for each study period was calculated.

I used nonmetric multidimensional scaling (NMS) to determine how the species presences changed in environmental space from the 1950s to the 1990s. NMS is an ordination technique that iteratively fits the data into a best fit ordination space, which is constructed from the multivariate environmental data (Minchin 1987). NMS uses a step-down procedure to find the most appropriate ordination dimensionality. The data points

are systematically moved around the ordination space until a minimum stress value, a measure of distance between the points in environmental and ordination space, is achieved. NMS ordination was conducted in the "vegan" in R (Dixon 2003, Oksanen *et al.* 2011).

Case Study One: Results

A total of 886 within-30-day blue-winged teal were used for the 1950s and 1241 for the 1990s. For the 1950s, the peak in fall velocity (i.e. the peak of the fall migration period) occurred in late October through the middle of November (Figure 8a). This peak shifted to early-to-late September in the 1990s. The beginning of the fall migration advanced approximately nine weeks in 1990s from the 1950s. Recoveries of banded blue-winged teal were being collected throughout the fall migration period in both the 1950s and the 1990s (Figure 8b).

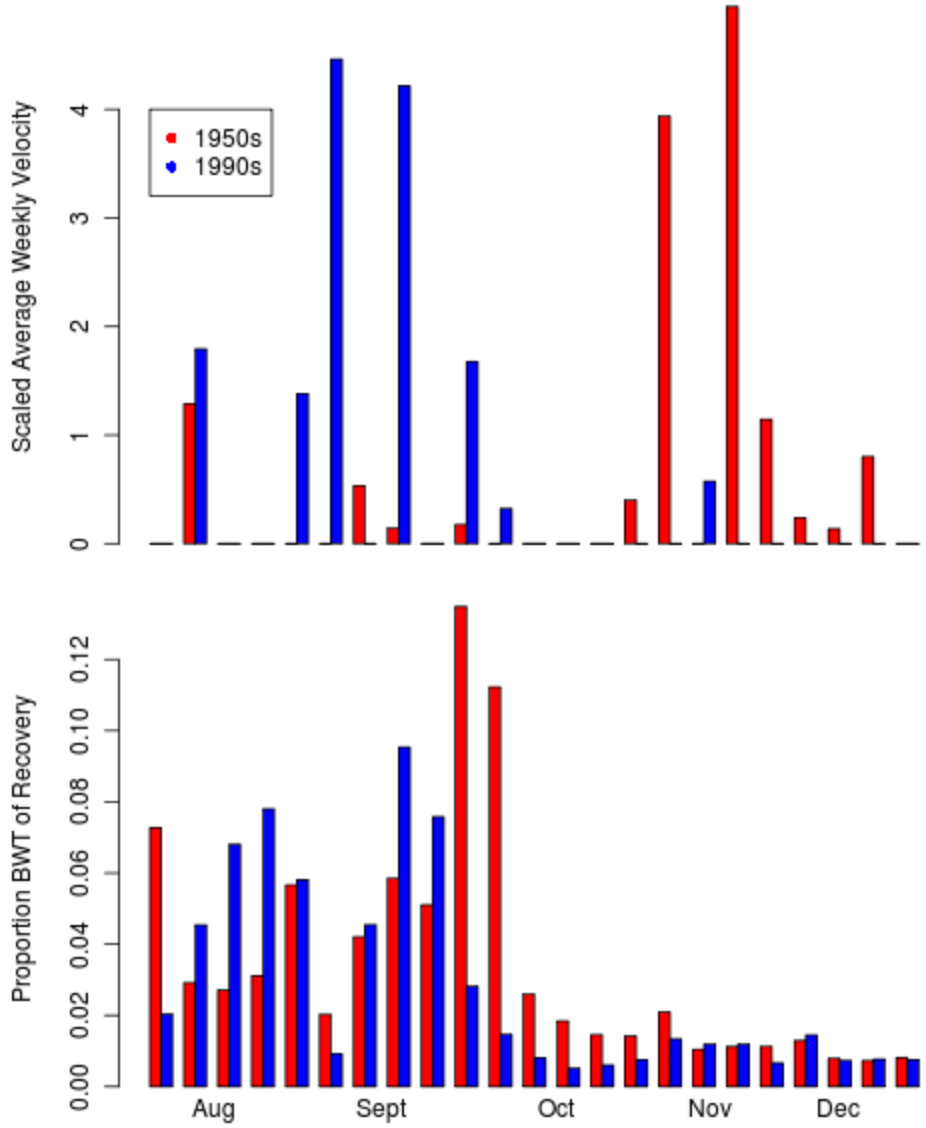

Figure 8A: Histogram of average weekly velocity.
The blue bars indicated the average weekly velocity of blue-winged teal recapture within 30-days of being banded for the 1990s, and red for the 1950s.
Figure 8B: Histogram of blue-winged teal proportion of the recovered BBL bands.
As in Figure 8A, the red bars are the 1950s data and blue bars are 1990s data.

In addition to exploring how blue-winged teal changed in spatial distributions, I examined how the MaxEnt models varied its predictions of probable occurrence relative to the environmental variables (Figure 9). I focused on temperature and depth to water table, each of these two variables has the highest AUC scores of MaxEnt models built iteratively with singular predictor variables. For 1950s fall, the temperature-only model has an AUC of 0.645 and the DWT-only model had an AUC of 0.689. The 1990s fall was 0.624 for the temperature-only model, and 0.745 for the DWT-only model. Even though the kappa statistic for the blue-winged teal fall is low (0.399) and the spatial location changes considerably (Figure 7), the overall response of the predicted probability of occurrence to temperature and DWT varies slightly between the two time periods (Figure 9). The 1950s have a peak in predicted occurrence around 6°C, and slightly higher in the 1990s (approximately 7°C). Blue-winged teal, in both study periods, have the highest predicted probability of occurrence below 20m DWT.

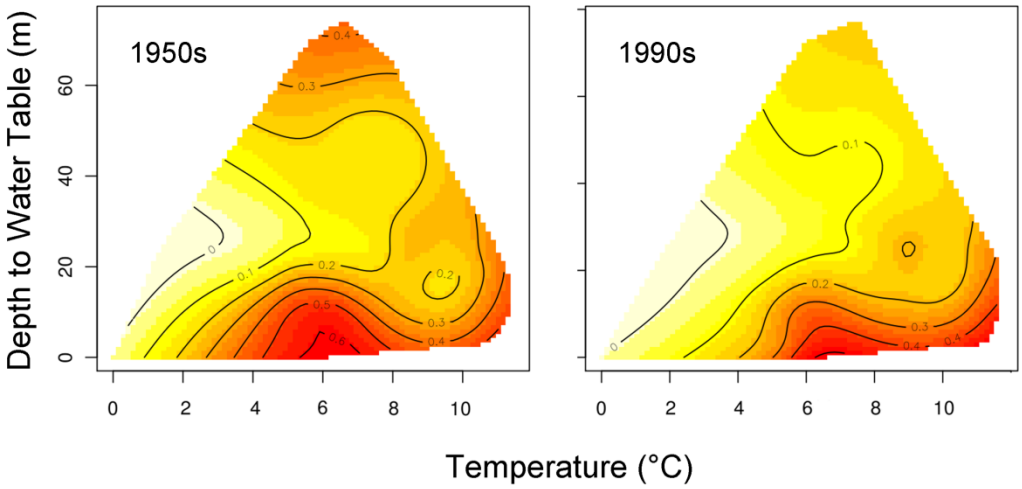

Figure 9: MaxEnt predicted probability response surfaces for fall blue-winged teal. The surface color ranges from yellow to red as it increases from low to high probability of occurrences. The surfaces are interpolated from 1,000 random points across the study area, and fitted contours.

Figure 10 depicts the results from the NMS ordination of the 1950s and the 1990s environmental data. The lowest stress value fit, 8.327, with two dimensions was well within acceptable range of model performance (Clarke 1993). The main result from this analysis is that the 1950s and 1990s points are completely intermixed along both of the NMS axes. In other words, there is no clustering of 1950s and 1990s into separate regions of this multidimensional ordination space. Based on these figures, it possible to conclude that there is niche conservation between these two time steps for this species.

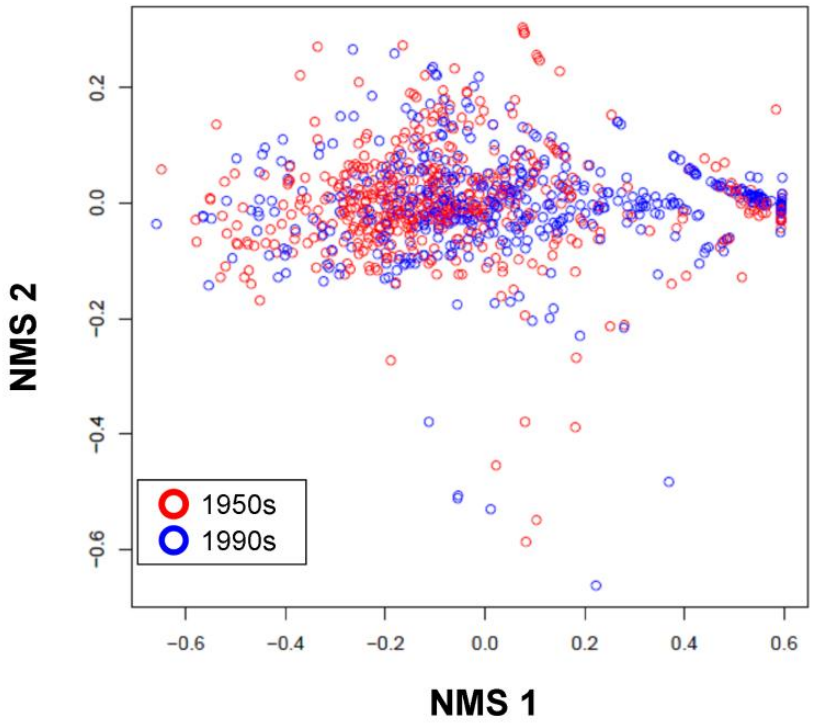

Figure 10: Nonmetric Multidimensional Scaling (NMS) ordination plot of blue-winged teal fall environmental data.
Red circles depict 1950s data point locations in the ordination space, and blue circles 1990s.

CASE STUDY TWO: LARGE SCALE RANGE EXPANSION OF WOOD DUCK

The large spatial changes in predicted probability of occurrence for wood duck in the fall from the 1950s to the 1990s are perhaps visually the most striking (Figure 7). In fact, wood duck has large areas of increased predicted probability in all seasons (Figure B1-B3). The low consistency between the two time periods is also reflected in the kappa

statistic, which ranges from 0.598 in the fall to 0.100 in the spring. During the gap between the study periods, wood ducks received increased protection from hunting pressure. I suspect that the increased conservation of wood duck is responsible for increase in the sample sizes for each period (Table 3). It was hypothesized that perhaps these rebounding populations were increasing their geographical and climatic ranges. First, I wanted to rule out that the increase in spatial distribution is due to sampling effort between the 1950s and 1990s. Then, I looked at the changes in predicted probability of occurrence relative to temperature and DWT.

Case Study Two: Methods

All 1950s BBL recovery locations were used to establish potential recapture locations during this decade. It was determined how many 1990s wood duck recoveries occurred in areas that were active recovery sites in the 1950s. If 1990s recoveries are predominantly located in areas where there was no recapture activity in the 1950s, then the argument could be made that the wood duck were present in these areas, but the areas were not being sampled. If 1990s recoveries are situated in areas where there was recapture activity in the 1950s, then these locations would be pseudo-absences for wood duck. Then it could be concluded that wood duck had increased its spatial distribution from the 1950s to the 1990s.

Case Study Two: Results

Approximately 83% of the 1990s wood duck recoveries occurred in areas where 1950s recapture activity was being done (Figure 11). Therefore, it is reasonable to conclude that the 1990s wood duck were being found in areas that it had not been in the 1950s. The expansion in area of high predicted probability of occurrence in the 1990s compared to the 1950s is real and not due to sampling in novel areas.

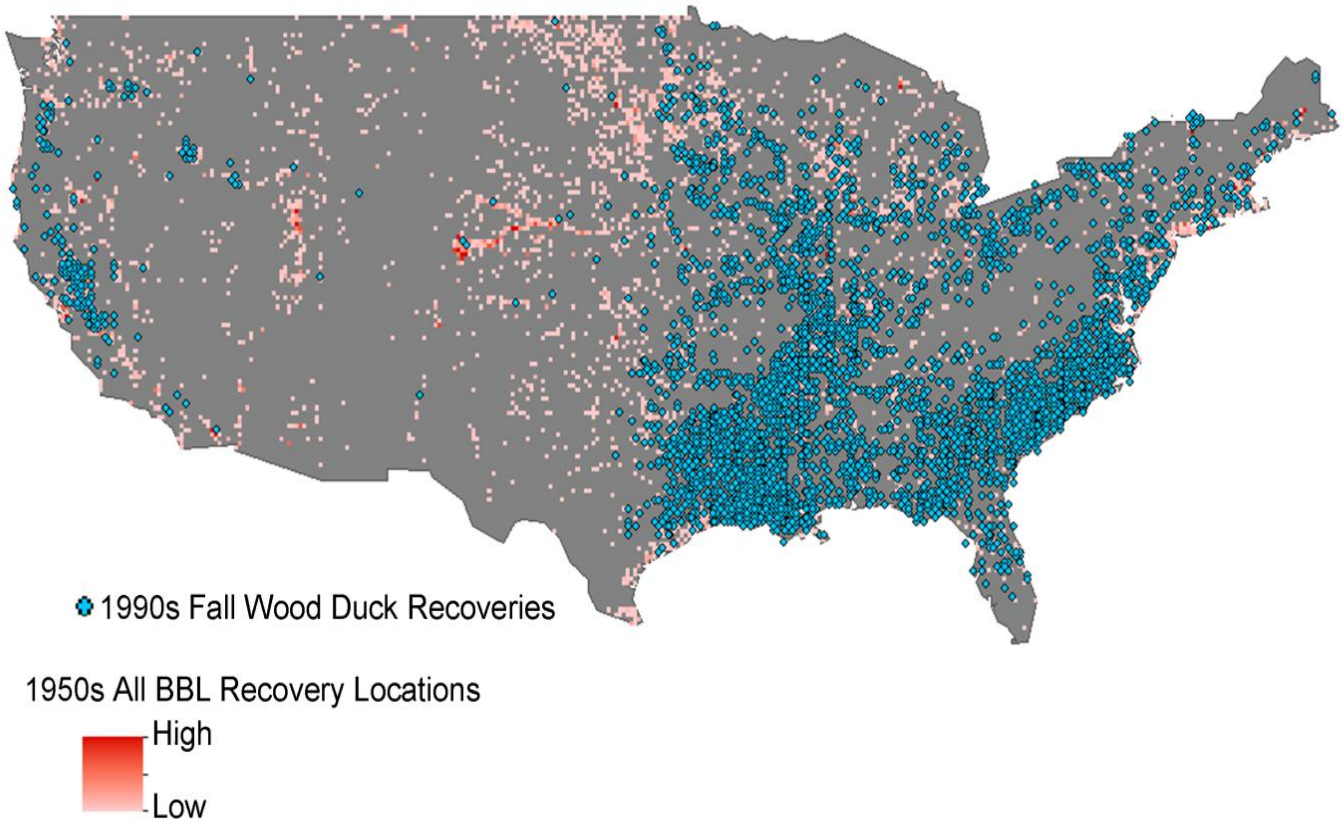

Figure 11: Map of 1950s fall wood duck pseudo-absences. The surface is coded according to intensity of 1950s recovery activity. The blue dots map 1990s recovery locations.

As with case study 1, the responses of predicted probability of occurrence to temperature and DWT were explored (Figure 12). For both time periods, the areas of high probability of occurrence are centered on higher temperatures and low DWT. The difference between the two decades is the lower predicted probability of occurrences. Wood duck has higher probability, even though slight, of occurring in areas with colder temperatures and high depth to water table.

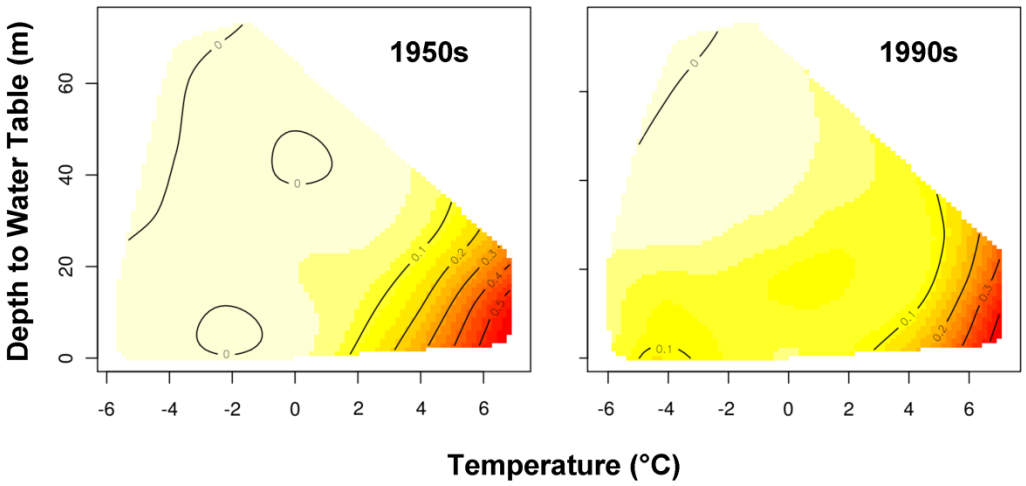

Figure 12: MaxEnt predicted probability response surfaces for fall wood duck. See Figure 9 for description and methods.

CASE STUDY THREE: SMALL SCALE RANGE CONTRACTION OF WOOD DUCK IN TEXAS

Even though wood duck had vast areas of increased predicted probability of occurrence in the 1990s compared to the 1950s, there were isolated pockets that had decreased predicted probability of occurrence (Figure 7). For this case study, the area along the Texas Gulf coast is the focus. This case study will attempt to address, when nearly all other portions of the United States had increase probability, why did this small portion of Texas have relatively large decreases in predicted probability.

Case Study Three: Methods

To determine why this area behaved differently than the vast majority of the United States, the values of the environmental variables were compared. Across the entire United States, approximately 1,500 randomly distributed points to determine the median and distributions of temperature, precipitation, and depth to water table were used. Within the Texan focal area (approximately 13 counties), approximately 200 random points were used.

Case Study Three: Results

For both the 1950s and the 1990s, Figure 13 illustrates the trends in temperature, precipitation, and DWT for entire United States and the area within Texas. Texas' temperatures, in both decades, are well within the extreme high temperature range of the

entire United States. However, the temperature between the 1950s and 1990s remain fairly consistent. The same is true of depth to water table. The values for Texas are within the lowest quartile of the values for the entire United States, but not different between the study periods. Precipitation for Texas is fairly comparable with the entire United States. However, Texas precipitation in the 1990s is higher than it was in the 1950s. Texas experienced a severe drought in the 1950s.

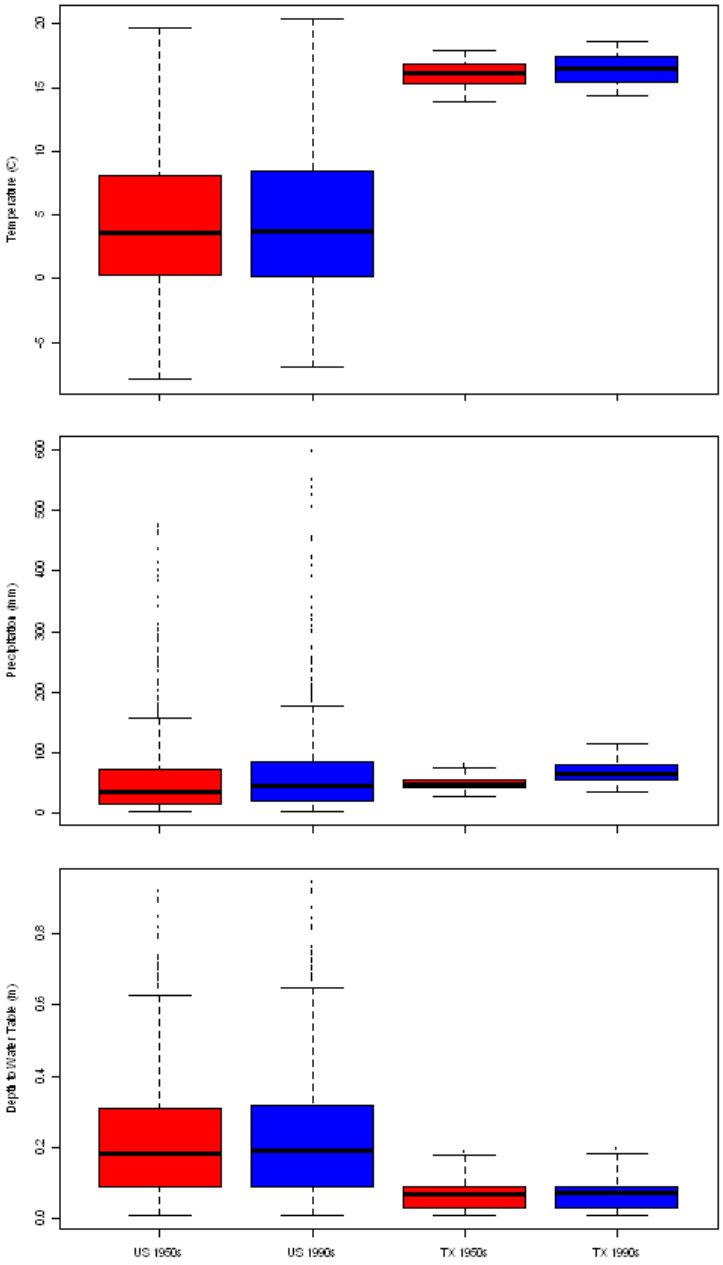

Figure 13: Box and whisker plots of temperature, precipitation, and DWT for fall wood duck. The red boxes are 1950s data and blue are 1990s. The left two boxes for each variable are from entire United States study area, and the right two plots are from the area of decreased predicted probability of occurrence within Texas.

DISCUSSION

Research, which relies on species distribution modeling as an analysis tool, often makes the assumption that distributions are static in time. Most commonly, research that attempts to make predictions about the future of species ranges has to assume that species range at the time of study is in equilibrium. This assumption is often false, however generally necessary due to the data limitations required to make inferences about changes in distribution through time. The goal of this research is to determine the relative stability through time of the spatial distributions for six migratory waterfowl species. By knowing how species respond spatially to changes in climate, we gain key ecological insight and increase our ability to manage these species long-term into the future. By using different species, we can compare how different species respond to varying environments.

Even though this research examines how species distributions change over time, it is not to be interpreted as an analysis of climate change impacts on species distributions. The study area is the contiguous United States, which does not encompass the entire range of the study species. Therefore, it is not possible to know if the range boundaries are changing in a manner that would be predicted by climate change (i.e. advancing poleward) (Parmesan and Yohe 2003). During the winter, the northern range boundary for some species is along the southern portion of the United States. There is some movement of this boundary northwards for species like the northern pintail (Appendix

B). To say that this is signal of climate change is impossible based on the work presented here. For one, this research only spans forty years, which is unlikely to be enough time to capture the response to anthropogenic climate change.

What this research clearly illustrates is that all species, even though very similar, respond uniquely to the changes in the environment. In fact, they all respond differently from one another and differently between seasons. All six of the study species are migratory waterfowl and share a lot of similarities. The differences between them, however, are still vast. Each species has an optimal temperature tolerance and ideal habitat within a specific depth to water table zone. These physiological constraints are compounded by factors such as behavior, community dynamics, and human influences. It was attempted to explore some of the factors influencing the distribution changes through the use of the case studies. The case studies are meant to give more awareness about the potential drivers influencing species distribution stability through time.

In the first case study, I further investigated the fall blue-winged teal pattern of change. Fall of the 1990s, relative to fall of the 1950s, had large increase in predicted probability in the southern portion of the United States, and large decreases in the northern portions (Figure 7). It was hypothesized that this change might be due to a shift to early fall migration. In other words, in the 1990s more blue-winged teal were already in the winter ranges during this period classified as fall. This is indeed what was found when comparing peaks in average fall velocity (Figure 8A). At first, these results seem to be counter to what we would expect according to climate change predictions. If

species are responding to warming trends, we would expect an advance spring migration and delayed fall migration (Schneider 1993; Peters and Lovejoy 1994). However, blue-winged teal is typically an early fall migrator; leaving for winter grounds as soon as their young have fledged (Bellrose 1976; Bent 1987). The results lead to us to question if the young of blue-winged teal are maturing more rapidly in the 1990s relative to the 1950s. And if their young are maturing more rapidly, does this mean they are leaving their breeding grounds earlier for their preferred habitat in the south?

Case study 2 looks are the large scale pattern of increased predicted probability of occurrence from the 1950s to the 1990s for fall wood duck. Actually, wood duck has large areas of increased probability for all portions of its annual cycle. It was shown that this increase in area is real, and not due to increased spatial extent of sampling. Also, it was illustrated how the predicted probability of occurrence is increasing into fringe environments. In other words, wood duck is more likely to accept colder habitat in the 1990s than it did in the 1950s. Wood duck is an interesting conservation success story. Previously, this species was intensively hunted due to the beautiful plumage of the male. In response to declining population number, hunting limits for wood ducks were drastically reduced. This rebounding population is one possible explanation for the large increase in BBL sample numbers from the 1950s to the 1990s. It could also be the explanation for increases in spatial extent. As populations increase in the core habitat areas, individuals of population might move into more fringe-type habitat. It is also

possible that wood duck are more capable of surviving in this area due to human influences, such as bird feeders and altered landscapes.

Wood duck predicted probability of occurrence increased throughout the United States in the 1990s, but there were small areas of decreased probability. Case study 3 explored one of these small areas along the Gulf of Mexico coast in Texas. The environmental conditions in this area have not changed that drastically between the two time periods. Precipitation did increase in the 1990s compared to the 1950s, but that is true for all of Texas. What has changed dramatically in this region is the land use. Two major dam projects were completed in this area between the two study periods. The dam on Lake Corpus Christi was completed in 1958, and then in 1982 the Choke Canyon dam was completed. Since wood duck prefer wooded wetland areas, these dam projects would have dramatically reduced its habitat. There is, additionally, extensive rice farming in this region, which would mean riparian habitat conversion. Despite the fact that the distribution modeling conducted for this research did not explicitly include land use change, it is possible to conclude that the impacts of the land use change in this region is still captured in my model.

Species distribution modeling has becoming a very important tool to look at how species will respond to future scenarios. Before predictions about the future are made, it might first be necessary to look at how species have responded in the past. All six study species varied drastically between each other and between seasons. The case studies begin to illustrate that there are numerous drivers of changes, and we can assume there

are countless not addressed here. The next focus for this research area needs to be how do we take this information about past responses and apply it to making predictions about future responses. For example, if a species has remained fixed in spatial distribution regardless of the climatic conditions, does this mean that it is unlikely to relocate in the future? Clearly, perfect future predictions are impossible and not testable, but it may be possible to significantly improve those predictions. After all, the best predictor of the future is the past.

Chapter Four: Predictive Migratory Movement Model illustrated through a Blue-Winged Teal (*Anas discors*) Case Study[3]

ABSTRACT

Historically, the migration of birds has been poorly understood in comparison to other portions of the annual cycle. The main goal of this research is to present a novel approach, illustrated via a blue-winged teal (*Anas discors*) case study, to predict the movement of migratory birds. The presented process incorporates not only constraints on habitat, but also approximates the likely bearing and distance traveled from a specific starting location. The presented method allows for movement predictions to be made from under- or unsampled areas across large spatial scales. The USGS' Bird Banding Laboratory database was used as the source of banding and recovery locations. Specifically, the recovery locations from banding sites with multiple within-30-day recoveries were used to build core maximum entropy models. The ancillary variables are measures of not only the environment (temperature, precipitation, elevation, and depth to water table), but also bearing and distance. The core models were used to project probability of movement from starting locations that did not have sufficient species data for independent predictions. The final model for an unsampled area was based on an inverse-distance weighted averaged prediction from the three nearest core models. To illustrate this approach, three unsampled locations to probabilistically predict where

[3] A version of the material presented in this chapter is currently in review at *Global Ecology and Biogeography*

migratory blue-wing teals would stopover were selected. These three selected locations, despite having little or none blue-winged teal data, are assumed to have populations. For the blue-winged teal case study, 104 suitable locations were identified to generate core models. These locations ranged from 20 to 228 within-30-day recoveries, and all core models had AUC scores greater than 0.80. From the appropriate core models, three projected models (Ontario, Montana, and New York) were built. It is concluded, based on model performance assessment, that this approach to predicting migratory movement is well-grounded and provides a good first approximation of bird migratory movement.

INTRODUCTION

Despite the vast amount of research conducted on birds, migration is poorly understood relative to the other portions of migratory birds' annual cycle (Bairlein 2008). In 1979, Gauthreaux called for a modern synthesis of bird migration, which without, he argued, we would be incapable of properly identifying and monitoring critical stopover habitat or assessing population dynamics of threatened and endangered species. As our needs for understanding migration remain the same, or are perhaps even more critical, we still have yet to reach a full synthesis on bird migration (Hutto 1998; Hutto 2000). The primary goal of this research is to present a method that will advance our ability to model the migratory movement of birds based on habitat requirements, directionality, and distance. I hope that this method, in the future, will provide a means by which to explore connectivity between breeding and non-breeding habitat, to identify key stopover areas, and to forecast potential impacts of climate and land-use change on migration systems.

The lag in knowledge gained regarding the migration of birds relative to breeding and wintering ecology is largely due to the intrinsic difficulties associated with studying migration. Clearly there has been significant work in migration ecology in the last thirty years (for a review: Faaborg *et al.* 2010), and key advances are due to innovations in technology, i.e. radio/satellite telemetry, GPS, and isotope analysis. However, modern technologies remain expensive, and thus greatly reduce the feasible sample size. Some example studies using these methods include 12 geese (Fox, Glahder, and Walsh 2003) and 7 eagles fitted with satellite transmitters (Meyburg, Scheller, and Meyburg 1995),

and 10 swans (Newman *et al.* 2009) fitted with GPS data loggers. These technologies have provided valuable information about fine-scale movement of individual birds, but yet we lack the ability to precisely extrapolate this knowledge to an entire population or to individuals in different areas.

In contrast to migration studies with small sample sizes, research has been conducted where comprehensive surveys count migratory birds at a key stopover location. Perhaps the best example of an area with nearly comprehensive migratory bird surveys is the Prairie Pothole Region (PPR) (Earnst 1994; Leon and Smith 1999; Naugle *et al.* 2001). The PPR is critical habitat for waterfowl; millions of migratory birds use the PPR as stopover habitat, while other birds rely on it for nesting sites (Williams, Koneff, and Smith 1999). To date, the PPR has lost an estimated 50-90% of its wetlands (Tiner 1984; Dahl 1990; Dahl 2006). Conservation of migratory waterfowl, which rely on this area, requires extensive knowledge about the ecology of the PPR. However, it also requires an understanding of the entire migration process. We need to know from where these birds are coming, where they are going, and what ranges in distance and direction they are capable of traveling (Erwin 2002).

Adding to the difficulty of studying migratory birds is that the traditional methods used to study animal movement have been developed, and perform best, for terrestrial animals. For example, if researchers were interested in determining the most likely route for an animal to move from one location to another, they might consider using a least-cost path analysis (Adriaensen 2003; Driezen 2007). This approach evaluates the

landscape between the source and destination, and finds the route where the animal will incur the least cost (e.g. calories burned, predator exposure, etc.). Least cost path is a difficult method to apply to migratory birds, because they are not forced through unfavorable habitats; they may simply fly over or around undesirable areas. The problem becomes how to predict bird movement when the methods developed for studying animal movement have a terrestrial bias.

Recently, there have been a few well-constructed attempts to spatially model migratory bird movement, specifically to determine their migratory pathway. The Tankersley and Orvis (2003) study explored potential migration pathways of neotropical birds across the United States. They determined location of optimal stopover habitat, defining habitat requirements from previous research, and then established connections of stopovers based on a fixed distance and bearing. In 2008, Downs and Horner developed a network approach to investigate potential migration pathways, which attempts to find the optimal path (i.e. path with fewest stops) across the landscape. The nodes in the network are all delineated freshwater wetlands in the study extent, and, as with the 2003 study, the distance is predefined.

My research advances upon the foundation of the above studies by predefining neither distance nor any other predictor variable. This has significant advantages, since distance travelled by a bird will impact its habitat selection. Research has shown that migratory birds are likely to continue flying in search of better habitat, or vice versa, select less suitable habitat if distance travelled is great (Moore and Aborn 2000; Alerstam

2001). By not fixing distance travelled, my model is more likely to capture some of this interplay between distance and habitat.

To overcome many of the difficulties of studying the migration of birds and to advance the state of spatially modeling bird migration, I relied on the Bird Banding Laboratory (BBL) dataset. Over the last 100-plus years, the BBL has gathered extensive data on banding birds (Gustafson and Hildenbrand 1999). Migration research was a key focus of banders early in the history of banding birds. Through time, this data has become a key data source in monitoring bird populations and determining appropriate hunting limits. Since the BBL has a large sample size (over 3 million game bird bands recovered) and large spatial extent (United States, and, although less comprehensive, Mexico and Canada) (Buckley *et al.* 1998), I feel that this is an ideal data set to accomplish the goals of this work.

The main goal of this research is to present a novel method, which will provide a means to predict migration movement of birds. Predicted movement from one area to another is contingent on habitat selection criteria, and likely distance and bearing of travel. This approach offers considerable benefits to the area of migration study. The described approach allows for predictions to be made from any starting or stopover location. This includes areas that might have little or no sample data. Additionally, as habitat availability and climate changes, the proposed method will allow for predictions to be made about anticipated effects on migration routes.

METHODS

The main goal of this research is to develop a method by which to predict movement of a migratory bird from a source location to its next migration stopover (or to, perhaps, its final destination). This endeavor, essentially, has two main lines of analysis. The first is creating models to determine the predicted probabilities of movement from source areas that have been thoroughly sampled. These models will be referred to as the core models for the entire process. The second, and the real obstacle for this area of research, is predicting movement to a destination from source areas where there has been inadequate levels of sampling to reasonably predict migration movement. To accomplish my goal, the models developed for well sampled areas (core models) were used to predict for less-than-ideally-sampled areas (undersampled point prediction).

Development of Core Models from Well-Sampled Areas

As previously stated, the BBL database was used as the source of biological data. The BBL is a long-term data set, started in 1902, which records the banding location of a bird and all its subsequent recoveries (Gustafson and Hildenbrand 1999). Nearly all bird species in North America have been banded during the course of the BBL history. To illustrate the approach of this study, blue-wing teal (*Anas discors*) fall migration was used as the test case, although my method is applicable to any migratory bird with adequate recapture data. Blue-wing teal was selected based on its relative high levels and consistency of sampling.

Initially, the BBL data was filtered to include only birds that were recovered within 30 days of being banded. For the remaining portion of this text, "recovery" will reference only to individuals recaptured within 30 days of being banding. Selecting these recoveries was necessary to increase the likelihood that the individuals' travel was not confounded by multi-season or foraging movement (Kölzsch and Blasius 2008). Additionally, only birds that moved from their original location were included. This avoided including birds that may not have begun their migration. Since all incidents are recorded to a 10-minute (approximately 16km) grid, how many recoveries each grid cell had was calculated. All grid cells that have above 20 recoveries were considered appropriate locations to build core models.

The ancillary variables used to model the destination probability from the selected source location are temperature, precipitation, elevation, depth to water table (DWT), distance, and direction. All variables were resampled to the BBL's 10-minute grid spanning the extent of the contiguous United States, Mexico and Canada. At 30-second resolution (1km), the average temperature and precipitation from WorldClim was utilized (Cameron *et al.* 2005). The 3-second (90m) Shuttle Radar Topography Mission (SRTM) Elevation Data Set was used for elevation (Homer *et al*. 2004). The DWT data layer is a simulated data set that reliably predicts the location of natural wetlands by finding the long-term stable solution of the balance between the climate driven fluxes (precipitation and evapotranspiration) and geologic/topographic water fluxes (riverine and groundwater

movement) (Fan *et al.* 2007). The DWT has been shown to be an accurate predictor variable for migratory waterfowl habitat (Kreakie, Fan, and Keitt *in review*).

The above ancillary variables remained constant in all the models, whereas distance and bearing varied according to source location. Distance and bearing were both calculated in R (R Development Core Team 2011) using the "geosphere" package (Hijmans *et al.* 2010). To determine the distribution of distance traveled from a specific source location, great-circle distance was calculated between the source location and all potential destination cells (i.e. to all other 10-minute grid cells in the study area). This process was repeated for direction, but to calculate the bearing between source point and all potential destination cells. These two variables were created for each prediction made from a single source location.

The probability of migration to a destination location was modeled using a maximum entropy algorithm executed in MaxEnt version 3.3.2 (Dudik *et al*. 2004; Phillips *et al*. 2006). MaxEnt is typically used in ecology for creating species distribution models, and the desired results are not exceedingly different from that of a species distribution model. Instead of creating predictions based only on the distribution of habitat measures, my approach includes consideration of the distance and direction of appropriate habitat. In essence, MaxEnt "estimate(s) the target distribution by finding the distribution of maximum entropy (i.e., that is closest to uniform) subject to the constraint that the expected value of each feature under this estimated distribution matches its empirical average" (Phillips, Dudík, and Schapire 2004). MaxEnt was also selected over other

approaches due to several of its attributes. It requires presence-only data, therefore it is not necessary to have known absences or create artificial absences (Elith *et al.* 2006). MaxEnt has been shown to model accurately despite varying sample sizes (Wisz *et al.* 2008). Even though the BBL is a large data set, it was subsetted substantially to obtain desired data sets.

Development of Undersampled Point Prediction and the Final Prediction

To predict movement from an area that does not have adequate recovery numbers for independent modeling, I relied on projections from core models created in the previously described approach (Figure 14). For undersampled locations, the nearest, in geographical space, core model was located. This nearest core model was used to project predictions for the undersampled point. The undersampled point predictions were created by retaining all the environmental independent variables, except the distance and direction were changed to correspond to the new source location. This process was repeated a total of three times with predictions made from the three closest core model locations. The three predictions were then averaged together weighted by the inverse distance to the specified point, which was used as the final prediction.

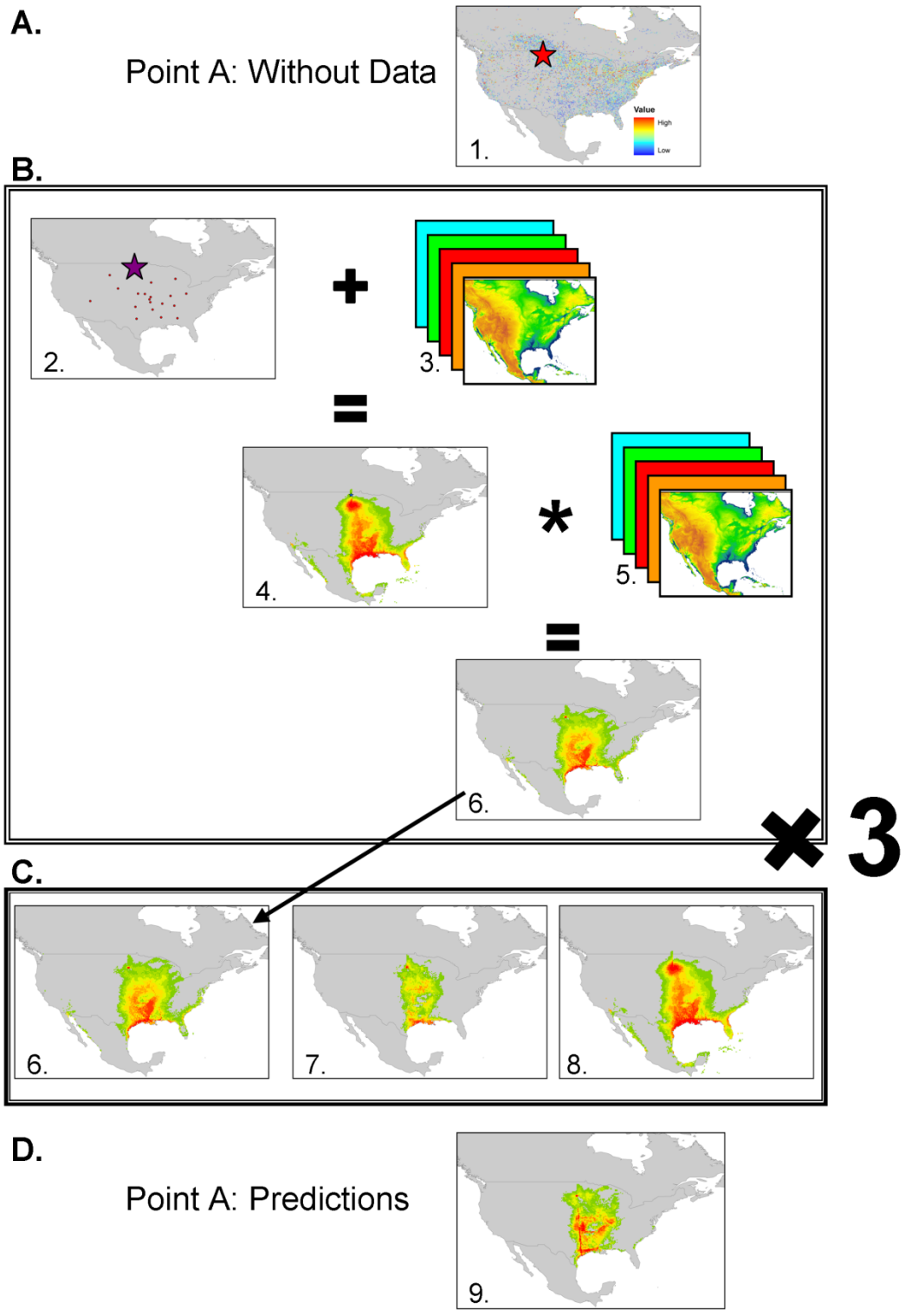
A.
Point A: Without Data
1.
Value
High
Low
B.
2.
3.
4.
5.
6.
3
C.
6.
7.
8.
D.
Point A: Predictions
9.

Figure 14: Graphical representation of the approach to predict movement from an undersampled point.

Part A: The point denoted with a red star is a hypothetical location where there is either no banding data or the available data is insufficient to create a core model. Part B: This portion is repeated three times; once for each of the closest three banding location with sufficient recoveries. Inset 2 depicts the location of the banding site (purple star) with multiple within-30-day recoveries (small red circles). The recoveries are used as the dependent variables with the independent variables (inset 3) to create a MaxEnt model of predicted movement (inset 4). With the ancillary variables (inset 5) recalculated for the undersampled point, the core model (used to create predictions in inset 4) is used to project the probabilities for the undersampled point (inset 6). Part C: The predictions from the first closest core model (inset 6) and the two next closest core models (inset 7 and 8) are averaged together weighted by the inverse distance to undersampled point. Part D: Inset 9 depicts the final averaged prediction of movement from undersampled point.

Assessment of Model Performance

The MaxEnt model for each of the core source locations was created by using only the within-30-day recovery locations as the dependent variable. Only the recoveries were used because we want to know where a bird is likely to migrate if it were to start at the source location. The recoveries were split into 70 percent training data and 30 percent test data, and the split was randomized ten times. A total of 10,000 background points, which were also randomized ten times, were used to produce the receiver operating characteristic (ROC) plot and calculate the area under the curve (AUC) score. The AUC score was used to determine acceptability of a model and to compare different models. Models with AUC scores over 0.7 are considered useful in predicting real world situations (Swets 1988).

As previously stated, the final prediction for an unsampled area is based on the average predictions of the three nearest core model locations. The decision to use three models was determined by comparing the performance of increasing numbers of core model projections averaged together to the predictions for a point with adequate within-30-day sample data. For clarification, first I selected a core model location and model movement predictions for this point (from now on referred to as the test location). Then I selected the next core model that was spatially the closest to the test location and projected movement predicted for the test location. The test location predictions and the projected predictions from the next closest core model were compared using a Kappa statistic. Kappa statistic measures the agreement between two data sets; a kappa statistic

of 1.0 is prefect agreement and 0 is no agreement (Landis and Koch 1977; Bell and Fielding 1997). Kappa statistic indicates how closely the forecasted predictions are to the prediction from actual data. This process was repeated by creating projected predictions from the next closest core model for the test location.

The above process for assessing the modeling methods for unsampled locations is not ideal, yet it is presently the only method available. Ideally, one would prefer to completely independent data to test the method, but this data is lacking. Therefore, I was required to use model prediction from the test location as comparison data set. Since I was required to compare model output to model output, this should be taken into account was reviewing the conclusions. However, the AUC scores of the core models were all well above the threshold of acceptable, and should be considered good points of reference. At least until independent movement data is available to validate the full method presented in Figure 14.

RESULTS

There were 104 locations that had over 20 within-30-day recoveries specifically for the case study on fall migration of the blue-winged teal. The number of recoveries for each core model location ranged from 20 to 228. Models preformed nearly equally regardless of number of recovery locations. The model with 228 recovery points had an AUC of 0.983, and the model with 21 recovery points, the least, had an AUC of 0.965 (Figure D1-A). Additionally, both the bearing and distance predictor variables significantly contributed to the AUC scores of the core model (Figure D1-B:C).

As a means of assessment for undersampled area predictions, I compared the outputs from a test location that had adequate sampling and then used my method to create projections for this area as if it did not have any sample data. The test location is in Thunder Bay, Ontario, Canada (48.42,-89.25), and had a total of 44 blue-winged teal birds recovered within 30 days of banding. The MaxEnt model was constructed with 44 recovery locations and had an AUC score of 0.962. All 103 other core model locations were used to project predictions for this test location. Kappa statistics, which describe the accuracy of the core model projected predictions to the predictions from actual data, ranged from 0 to 0.7978 (Figure 15A). Generally, the core models closest to the test location performed best. I tested the agreement between increasing numbers of core model projected predictions averaged together weighted by inversed distance to the test location (Figure 15B), and determined three to be optimal.

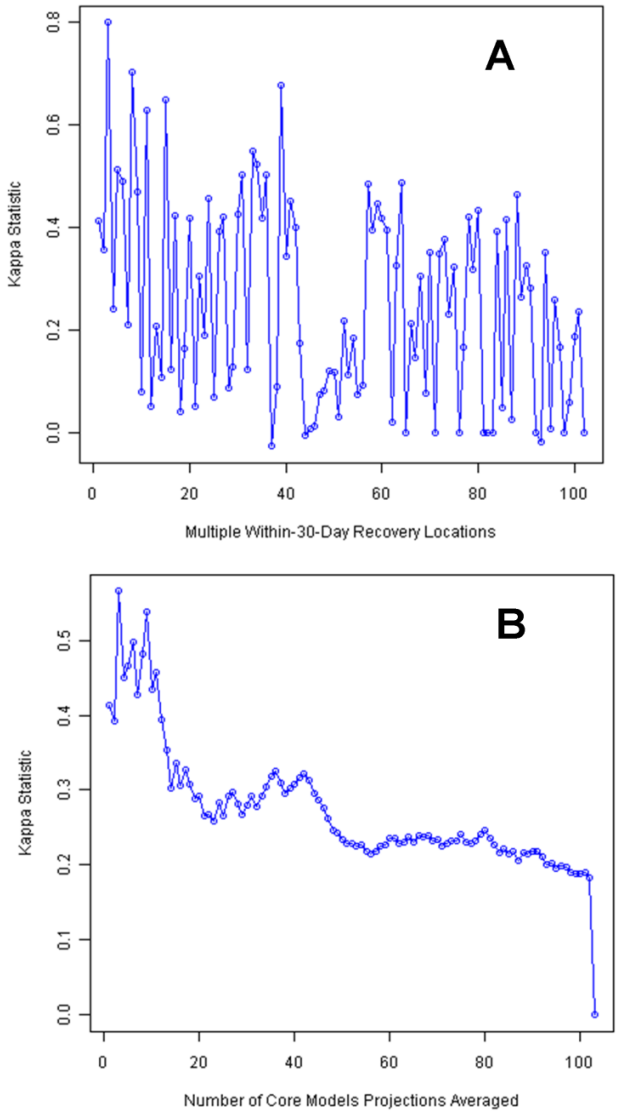

Figure 15: Line graphs of method assessment.
Figure 15A: Kappa statistic comparing the test location prediction to predictions made for the test location from the other 103 core model locations. The 103 core models are arranged with increasing geographical distance from the test location. Figure 15B: Kappa statistic comparing the test location to increasing numbers of averaged predictions made from core model.

Figure 16A maps the predictions for the test location based on actual within-30-day recovery data from this point. Figure 16B used the three closest core model locations: (1) eastern Montana, USA (48.48, -95.92) with 30 recovery locations, (2) Upper Peninsula, Michigan, USA (46.25, -85.92) with 30 locations, and (3) eastern Wisconsin, USA (44.08, -87. 92) with 23 locations.

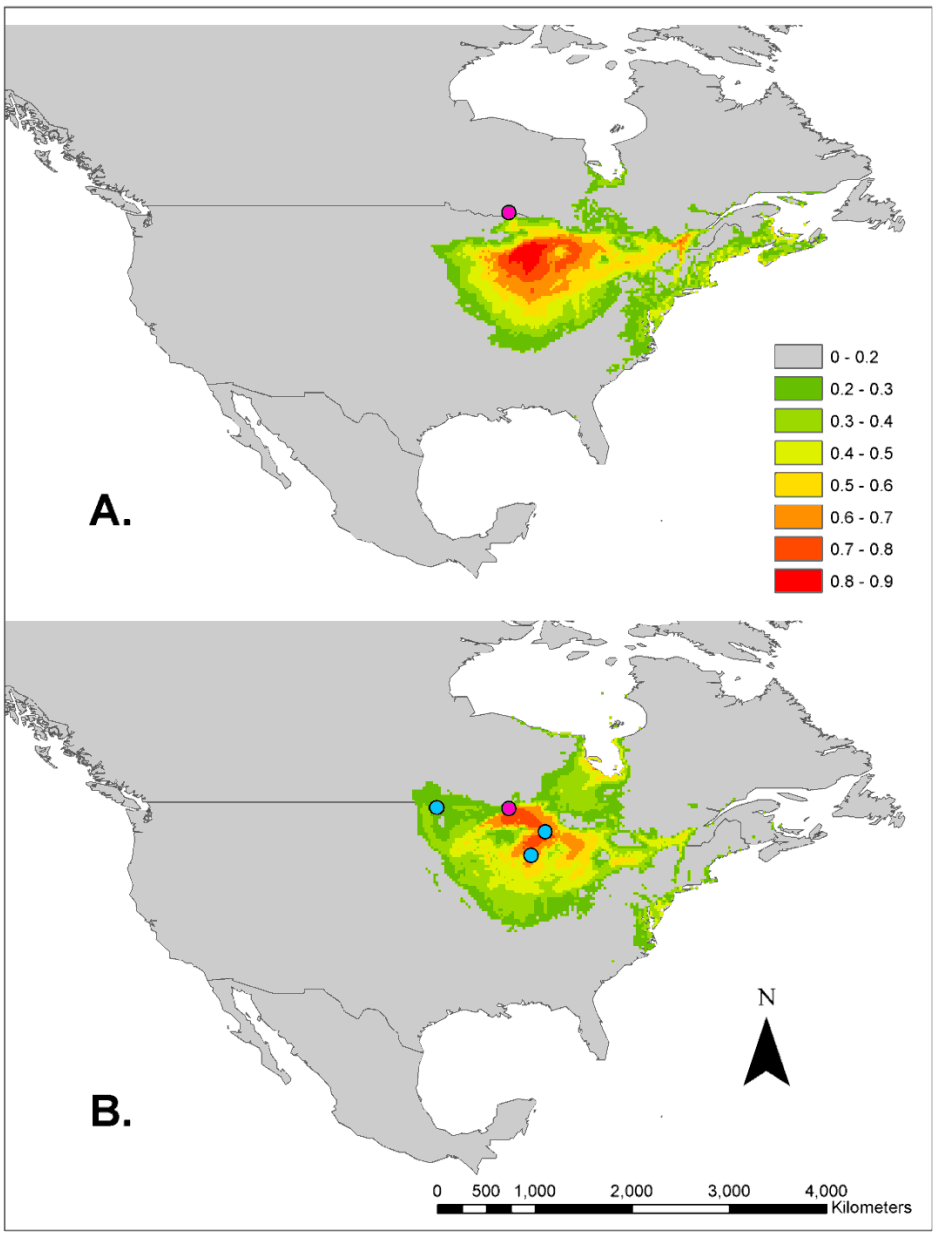

Figure 16: Maps of predictive blue-winged teal movement from Ontario, Canada (48.42,-89.25), which is marked with the red dot in each map.
Figure 16A is the predictive movement surface created with 44 within-30-day recovery locations. Figure 16B is the probability surface created using the method introduced in this text; essentially treating this point as if it had no survey data. Core models from the three location marked with blue dots were used to project for the point of interest in Ontario.

Additionally, I examined the predicted probability of occurrence for recovery points used to create the MaxEnt surface in Figure 16A (Figure 17). For the same recovery point locations, the values of predicted probability of occurrence were then compared to the values of projected predicted probability in Figure 16B. This provides us with a way to compare the predicted probability of occurrence for actual data points between the two methods presented in this research. We can see that the method that relies on averaging the three closest core models closely mirrors the model predictions created with actual recovery data. The projected predicted probability is however less than the predicted values from actual data, which is what we would anticipate. The method used to create predictions for location that lack sample data closely mirrors the predictions made by the model created with actual data, but is conservative in its predictions of likely occurrence.

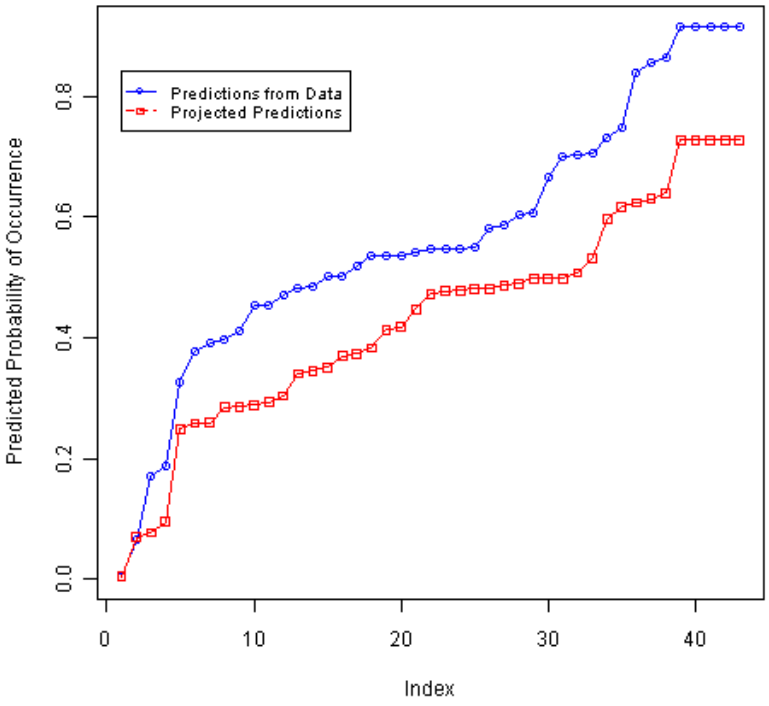

Figure 17: Line graphs of predicted probability of occurrence for recovery points.
The blue line represents the sorted values of predicted probability from the model created with actual date (Figure 16A). The red line represents the sorted value of predicted probability averaged from the three projected models for the test point (Figure 16B).

My method was then applied to three locations with no recovery data, and where it was considered highly likely to be blue-wing teal habitat (Kreakie, Fan, and Keitt in review). The first location was immediately south of Winnipeg in Manitoba (49.31, -97.34) (Figure 18A). This location's predictions are focused in the Mississippi flyway and have an upper probability of predicted occurrence around 0.72. The second model was built for a location in eastern Montana (48.37, -105.86) in the Prairie Pothole Region (Figure 18B). Its results, with highest predicted probability of occurrence about 0.71,

predicted a main corridor of migration through the Great Plains and eventually merging into the Mississippi flyway. The final location is in the Ha-De-Ron-Dah Wilderness Area in the New York Adirondacks (43.76, -75.21) (Figure 18C). This projection depicts less distance traveled and is concentrated in the Adirondacks and along the Atlantic Flyway. It projected 0.52 as the high end of its predicted probability of occurrence.

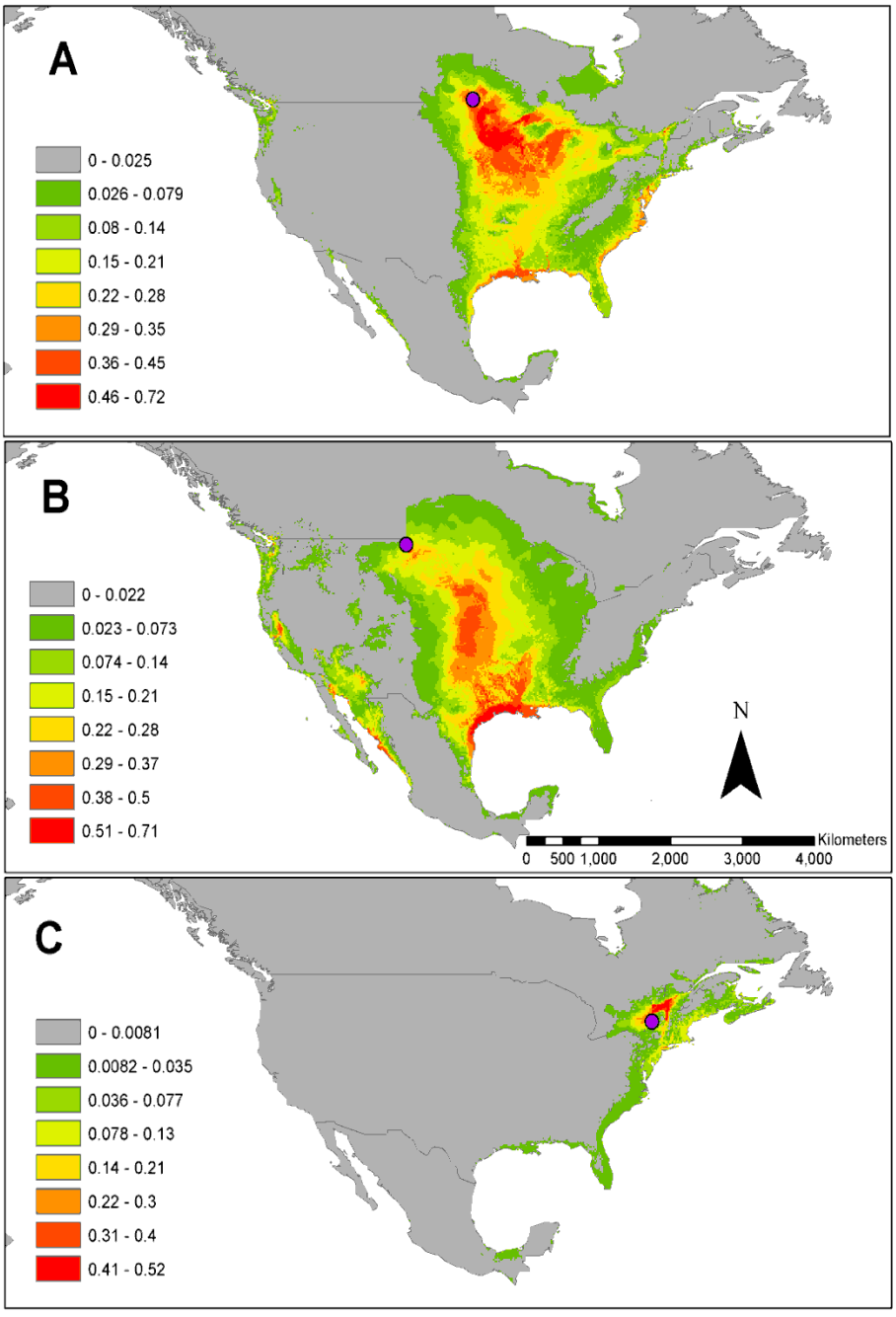

Figure 18: Maps of predictive blue-wing teal movement from three different source locations.
Figure 18A has a start location (purple dot) in Manitoba, Canada. Figure 18B is in Montana, USA, and Figure 18C in New York, USA.

DISCUSSION

The intent of this project is to present a novel method of predicting migratory bird movement and to explore its validity and outputs with a blue-winged teal case study. The inclusion of habitat measures, distance, and bearing make it possible to create probabilistic predictions based on the ecology, physiology, and behavior of a species. Each species is ecologically constrained by its suitable stopover habitat. Determining species distributions is, perhaps, the traditional manner in which MaxEnt is used in ecology; it defines the probability distribution of environmental variables. Beyond habitat needs of a species, there are physiological considerations that influence the optimal distance flown by birds between stopover locations (Weber *et al.* 1999). Clearly, migration has a predefined final destination, and this behavioral factor is accounted for by the inclusion of bearing.

Conceivably the largest contribution of this approach is the inclusion of both distance and habitat in determining the predicted movement. Given that migration is extremely costly to the animal in terms of resources and mortality, optimal migration theory illustrates that birds should attempt to accomplish migration with the minimal number of stopovers (Desholm 2003; Fujita *et al.* 2004). Since migration has a high caloric demand, physiology prevents individuals from making the journey in one step. This means that birds should attempt to maximize the distance they fly between stopovers while attempting to select optimal habitat for refueling and predator-avoidance. Stopover locations that are too close to one another will increase the overall cost of migration.

While flying past optimal habitat to maximize the distance between stopovers, the bird risks not finding suitable habitat before resources are depleted. The interaction between maximizing distance and optimal habitat selection during migration is directly incorporated in this method.

This method is fundamentally comprised of two main parts: (1) construction of the core models for banding sites that have adequate within-30-day recoveries and (2) using the core models to project probabilities for sites that lack sampling. Model assessment of each of these parts entails different approaches. Core model assessment relied on the use of AUC scores (Figure D1). Each core model performed very well, and can be interpreted as a reliable prediction of real-world events (Swets 1988). It was demonstrated that the addition of bearing and distance added to the accuracy of the predictions. Based on these results, it can be concluded that this approach for modeling migration movement from locations with numerous within-30-day recoveries is well founded. However, there are a great number of locations without the data necessary to directly predict movement, and must rely on the second part of my method.

The assessment of the predictions for undersampled location is not as forthright as the core model validation procedure. To illustrate the validation of undersampled location predictions, the predictions made from a core model as test data set were used. Then various projected predictions were created for the test location, and compared to the original predictions. The results presented in Figure 15 and 16 depict this validation

procedure. In Figure 15A, kappa statistics compare all core models projected predictions to the actual data predictions for the test location. Clearly, as the distance from the test location increase the kappa statistic decreases. For this point, projected predictions' accuracy decrease as distance between the core model location and projection point location increases. Instead of relying on a single projected prediction, I tested the validity of using multiple projected predictions for the final prediction. This agreement between the actual data predictions and averaged projected predictions spikes at three. These results show that the closest core models have the highest accuracy, and that three predictions averaged together is the strongest.

Although the predictions are not identical, and we have no way to know which is superior, they do share some key similarities. The areas in each map that have high predicted probability of being stopover habitat for this source location have the same spatial extent. This includes the complexity of the shape to the high probability area. For example, both results include areas around the Hudson Bay and up into the St. Lawrence River.

Even though the predictions are all made for the same species, this method makes it possible to examine regional differences in migrations routes due to the inclusion of distance and bearing in modeling (Figure 18). The maps for Manitoba (Figure 18A) and Montana (Figure 18B) depict similar predictions along the Central and Mississippi Flyways. However, those individuals leaving from Montana have a more diffuse predicted path, which includes some portions along the Pacific Flyway. This is obviously

due to the relative proximity to this area compared to the source location in Manitoba. The source location in New York (Figure 18C) is even further east, and nearly completely abandons the Mississippi Flyway in favor of the Atlantic. The ability to predict regional movement of migratory birds will be critical for their long-term conservation (Haig, Mehlman, and Oring 1998).

I relied heavily on the BBL data to implement this method. The BBL data has many attributes that contribute to usefulness in migration study: for example, the large spatial extent of banding sites, the near global recording of recoveries, and the long time span of the study. Nonetheless, there are data considerations that should be kept in mind while interpreting the output. Since the occurrence of banding sites that have sufficient within-30-day points to build a core is rare, they are not evenly distributed across the landscape. If the three core models used to create the averaged prediction for a site are all exceedingly distant from the prediction site, the results should be carefully reviewed. It is likely that the habitat selection would not change for migrating birds within a species. Although clearly the average distance and bearing traveled would vary regionally for migrating birds.

The need for increased understanding about migration is no less important than in 1979 when Gauthreaux called for a modern synthesis of bird migration. The field has advanced in the last 30 years, but not sufficiently to address his original concerns and definitely not enough to help mediate modern concerns. The method outlined in this text will allow for us to begin to explore step-wise movement across the landscape and

identify key stopover locations. Stopover habitat is of critical importance to migratory birds (Alerstam *et al.* 1990; Alerstam and Hedenström 1998; McWilliams *et al.* 2004). It has been estimated that some individuals spend as much of 90% of their migration actually resting and refueling in stopovers (Schaub *et al.* 2001). Typically, these habitats are wetlands, which are extremely sensitive landscape features (Fretwell *et al.* 1996). Wetlands are exposed to multiple anthropogenic pressures; such as water divergence, land use change, sedimentation/erosion, and rapid climate change. It is critical to understand how birds use these habitats as they move between their breeding and wintering ground to ensure their long-term conservation.

Appendix A: Chapter Two Supplemental Material

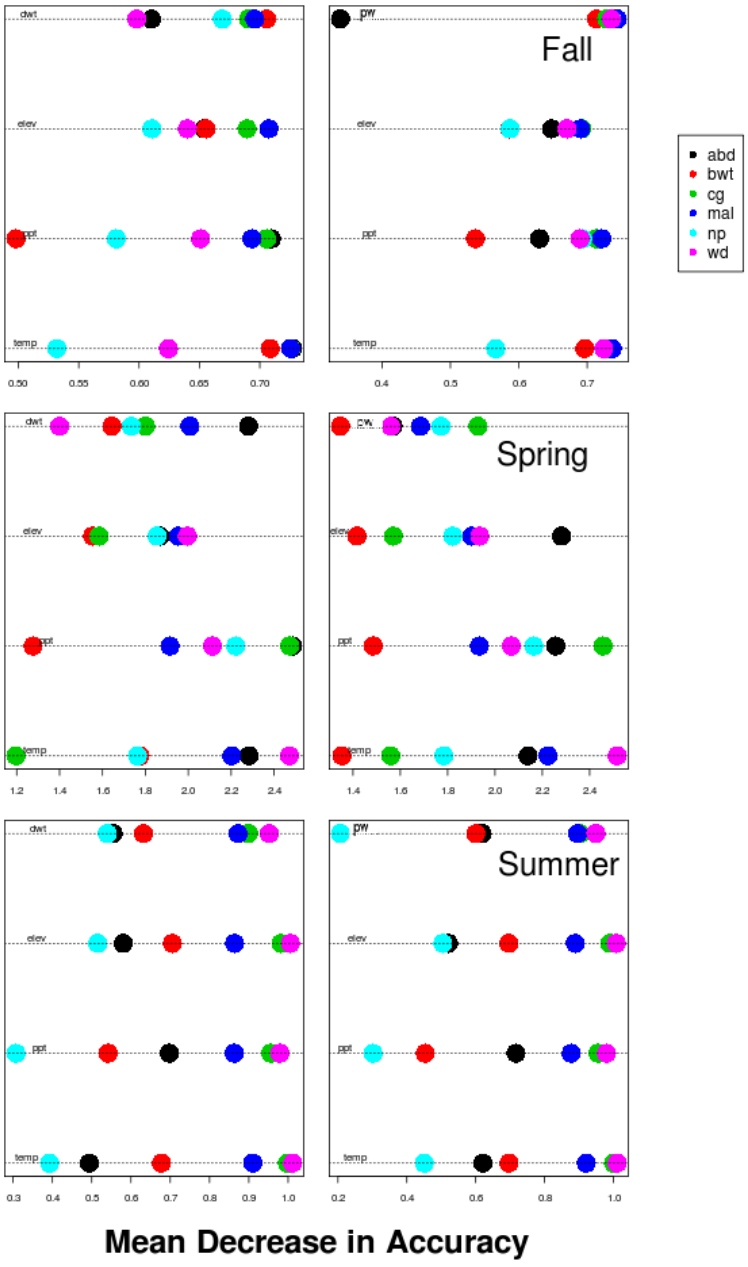

Figure A1: Plot of variable importance measure from Random Forest.
Variable importance is measured in mean decrease in accuracy, which is the decrease in accuracy of a classification after the variable has been randomly permuted. A higher mean decrease in accuracy means the variable contributes more to the accuracy of the classification. The abbreviations are as follows: abd (American black duck), bwt (blue-winged teal), cg (Canada goose), mal (mallard), np (northern pintail), wd (wood duck), temp (temperature), ppt (precipitation), elev (elevation), dwt (depth to water table), and pw (NLCD's percent wetland).

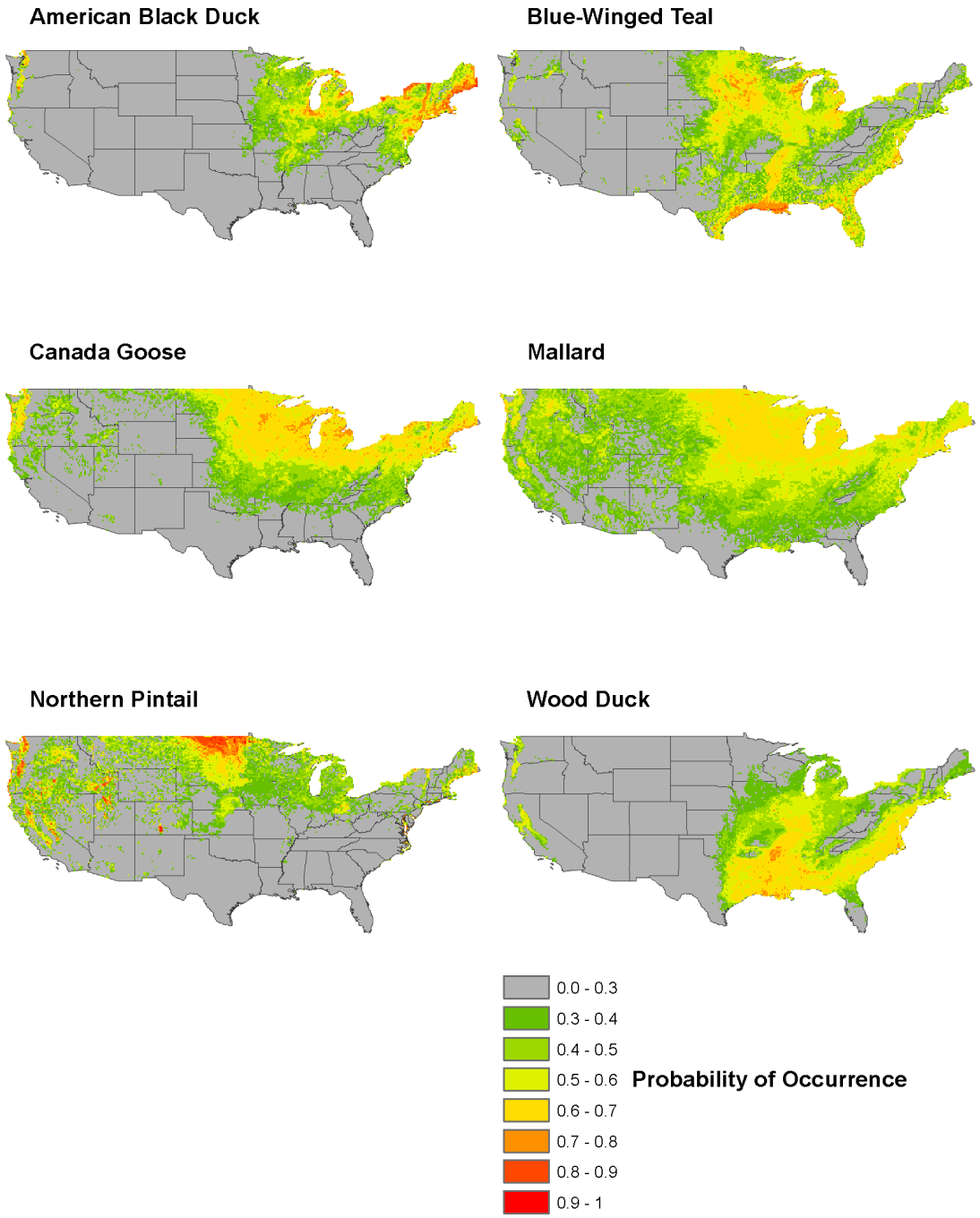

Figure A2: Maps of predicted probability of occurrence for all study species' fall habitat. Predictions were created using MaxEnt with 100% of known presence locations to increase accuracy of the visual representation. Temperature, precipitation, elevation, and water table depth were the predicted variables used to construct the probability surfaces.

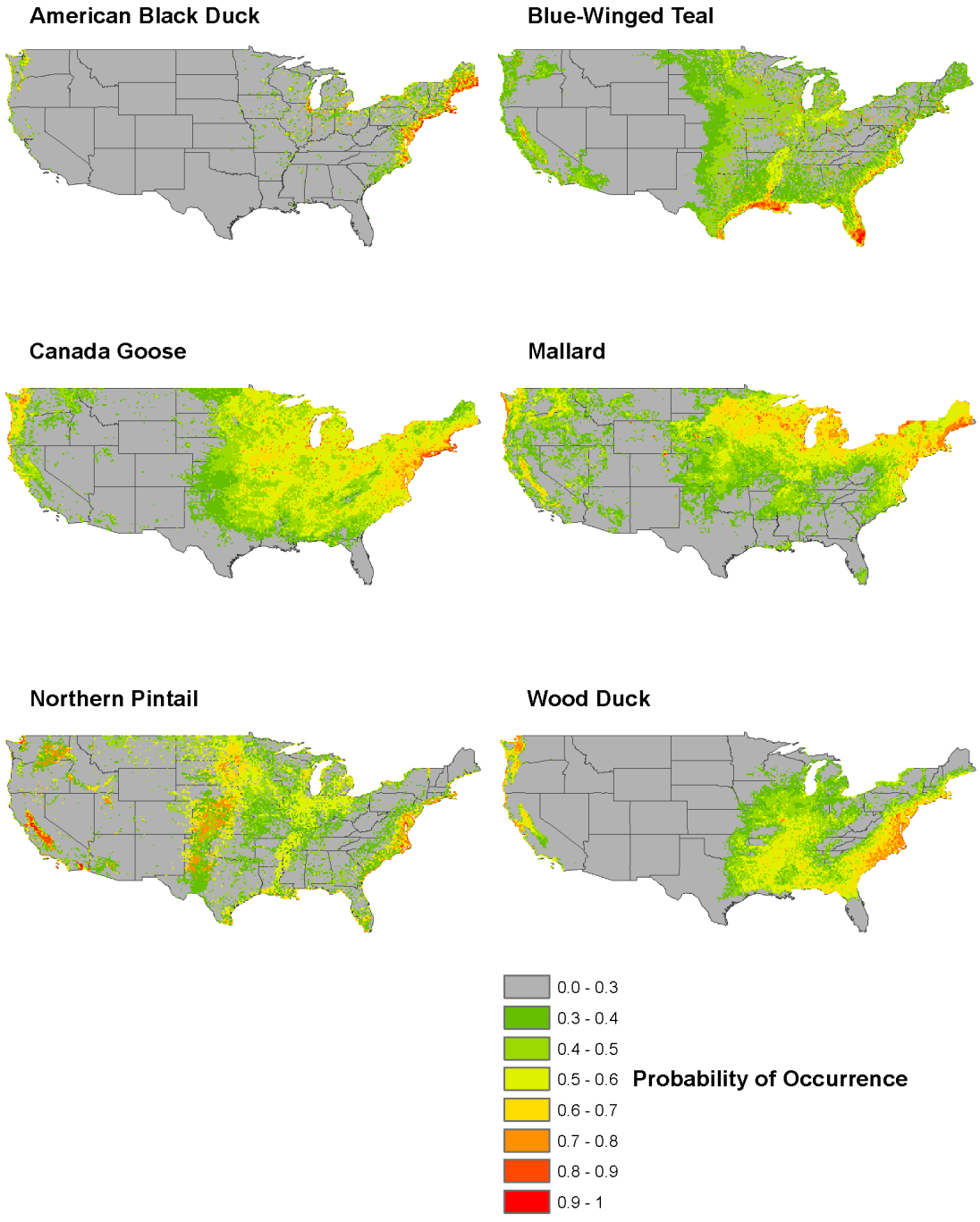

Figure A3: Maps of predicted probability of occurrence for all study species' spring habitat. See Figure A2 for description

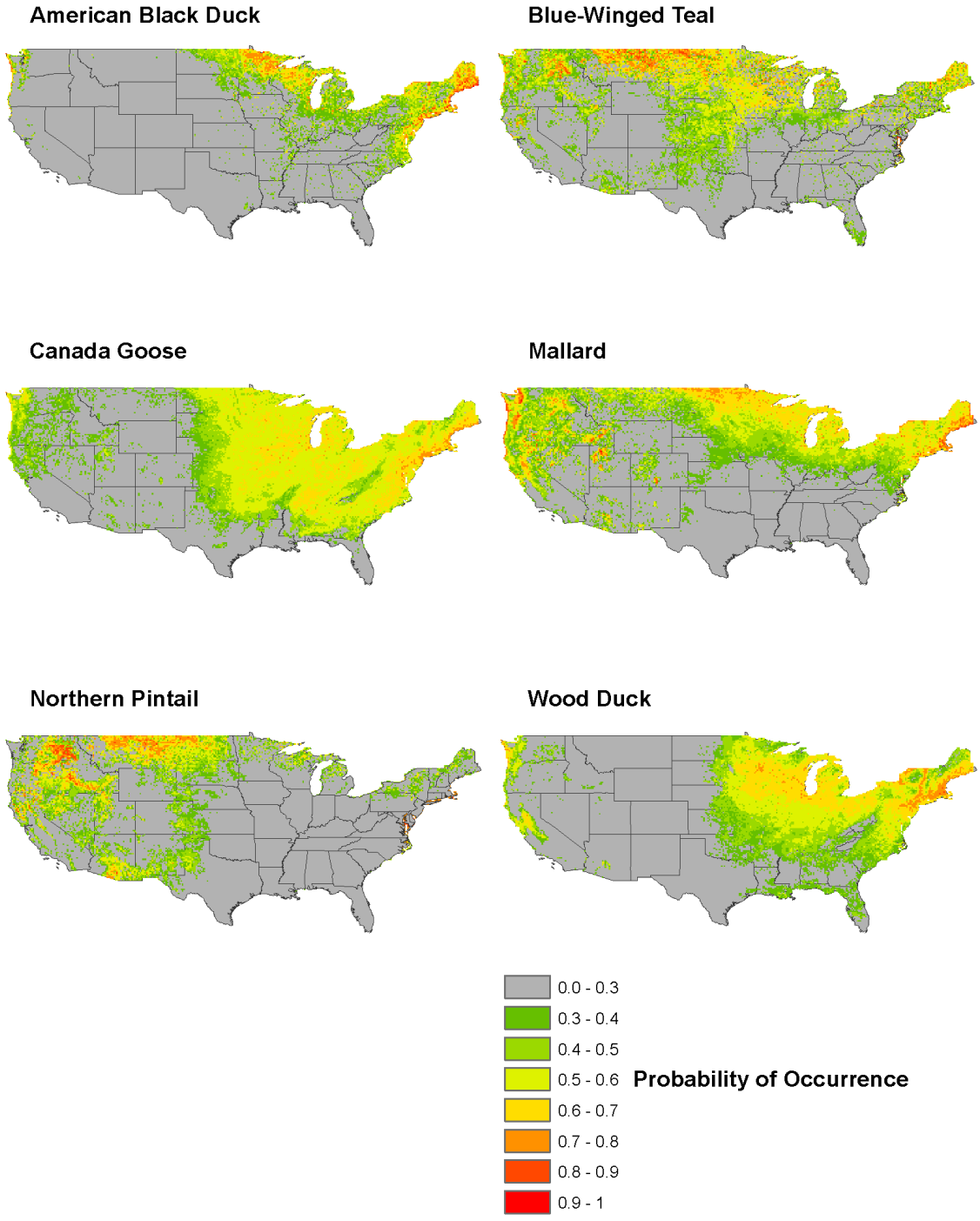

Figure A4: Maps of predicted probability of occurrence for all study species' summer habitat. See Figure A2 for description

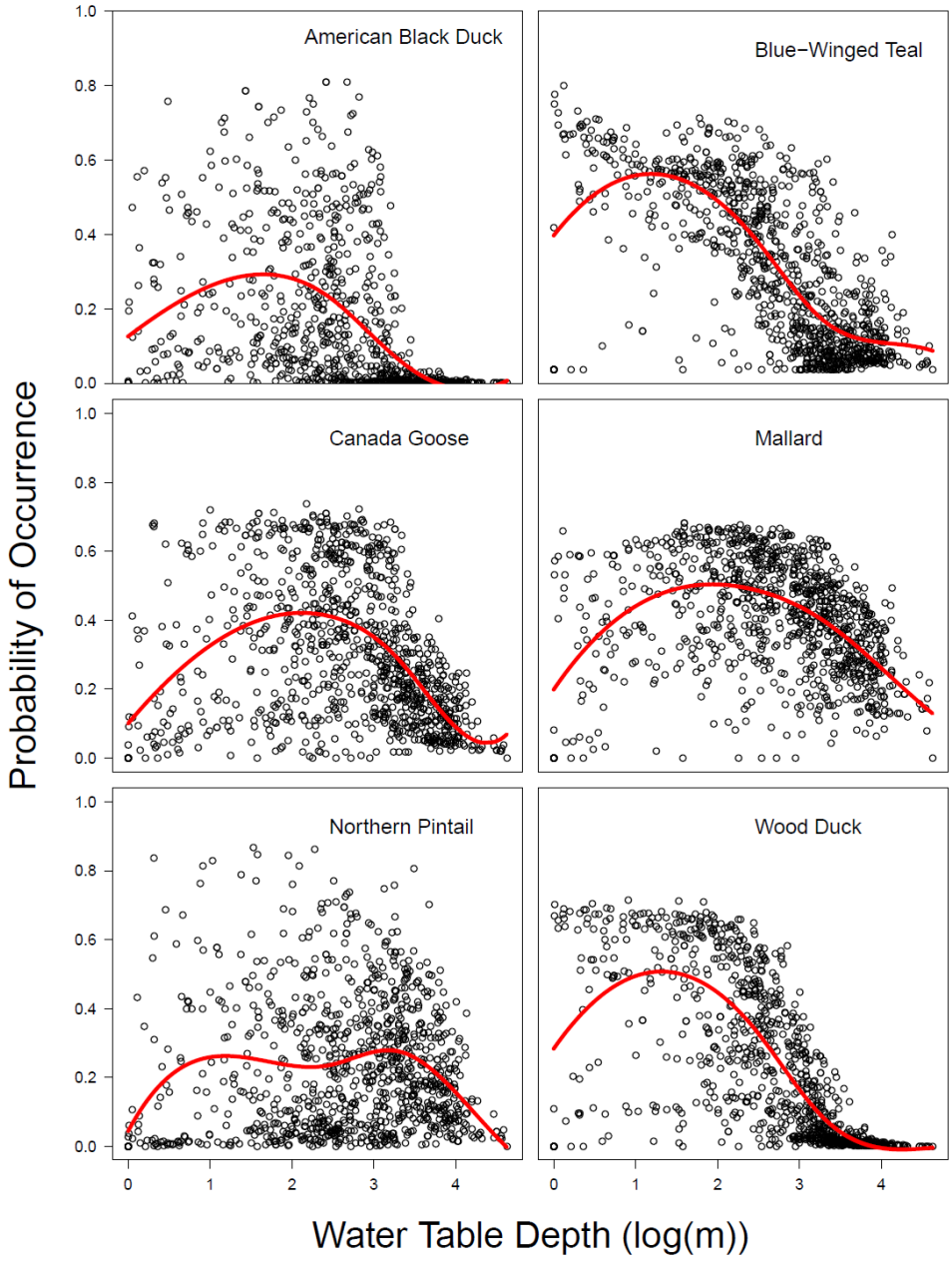

Figure A5: Plot of the relationship between the log transformed water table depth (m) and predicted probability of occurrence for each study species in fall.
The plots were constructed by selecting 1,000 random points from the predicted probability of occurrence surface. The red curve is a smoothing spline fit to the mean of the data points, and meant to illustrate the trend of the data.

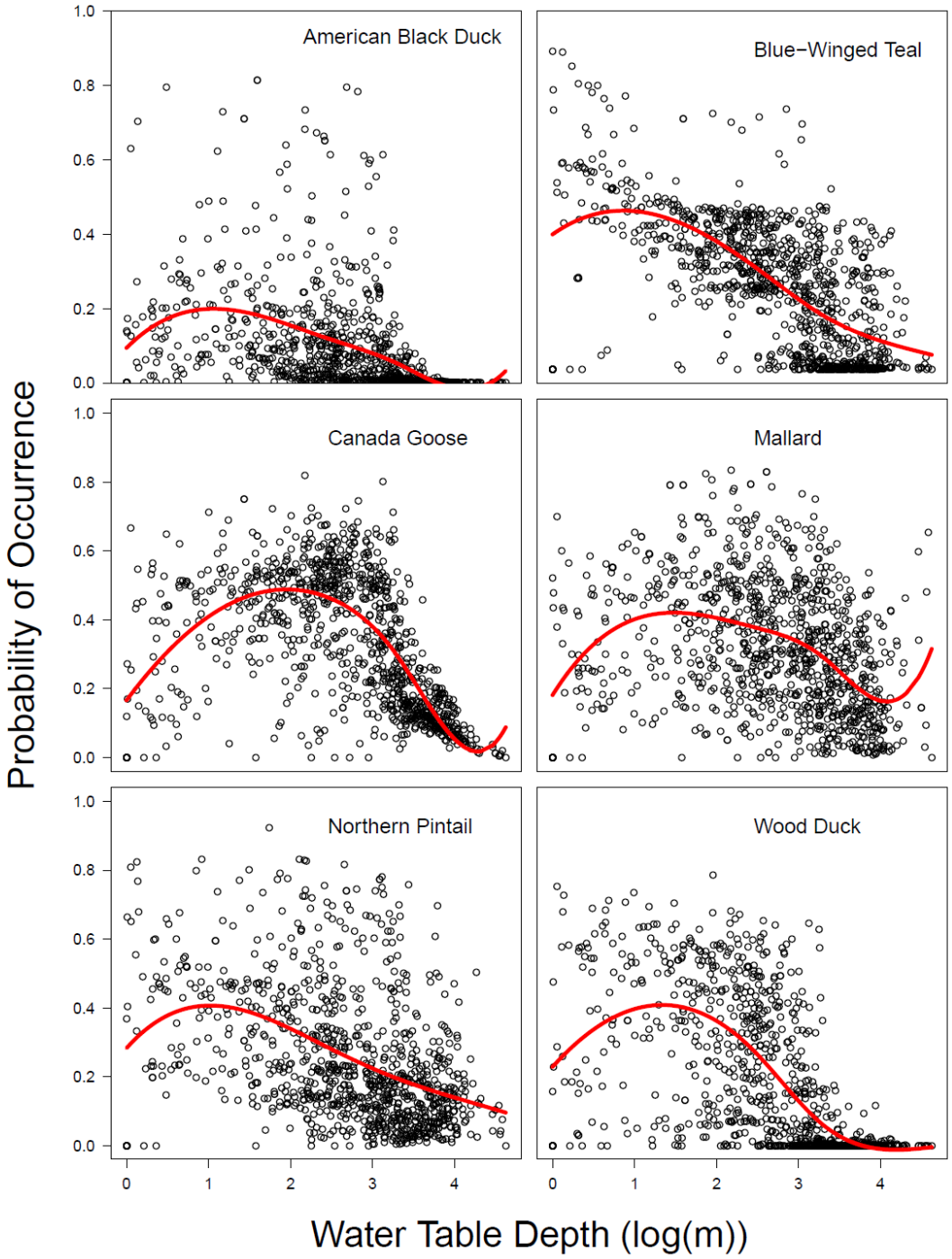

Figure A6: Plot of the relationship between the log transformed water table depth (m) and predicted probability of occurrence for each study species in spring.
See Figure A5 for description.

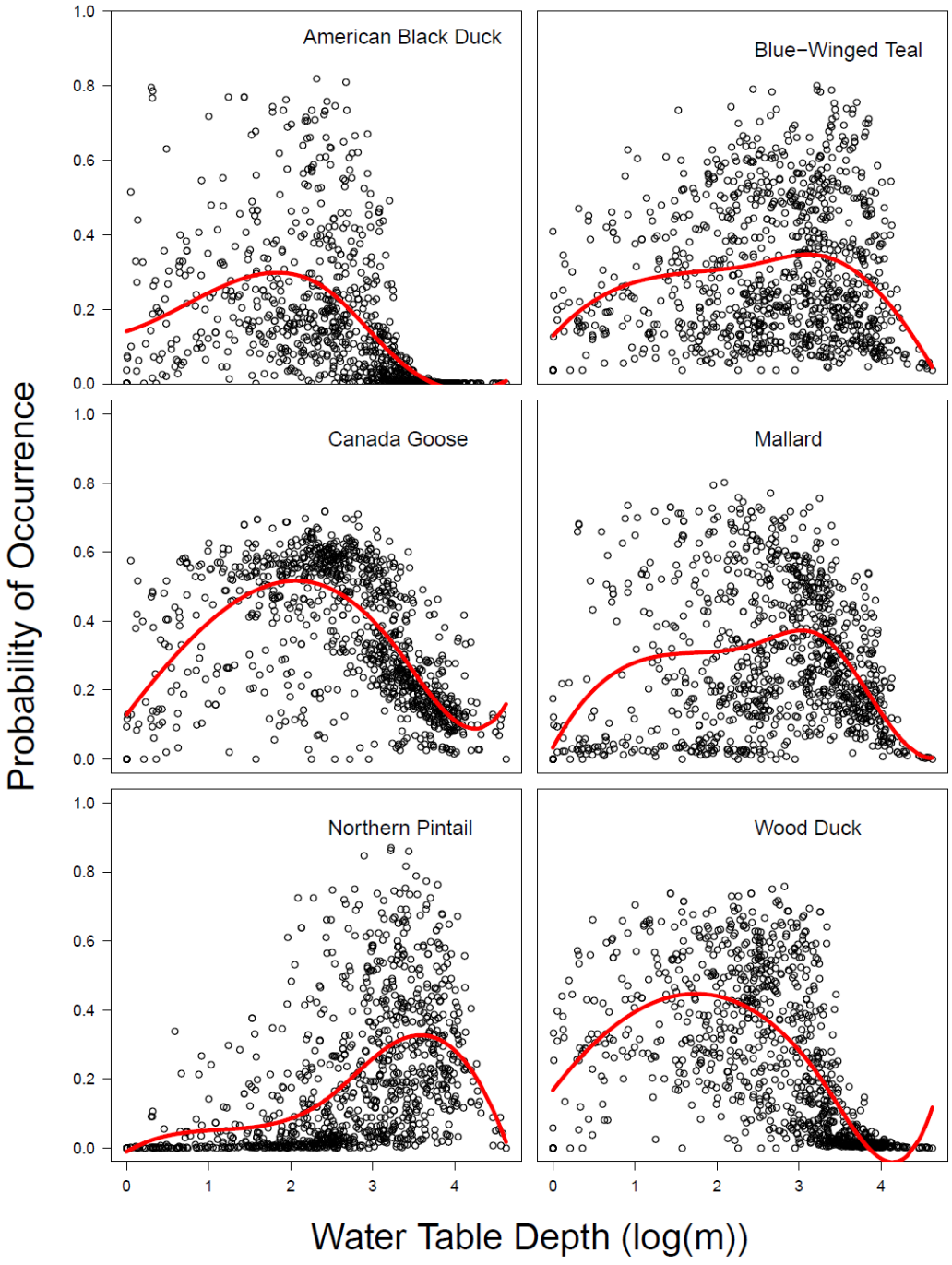

Figure A7: Plot of the relationship between the log transformed water table depth (m) and predicted probability of occurrence for each study species in summer.
See Figure A5 for description.

Appendix B: Chapter Three Supplemental Material Part One

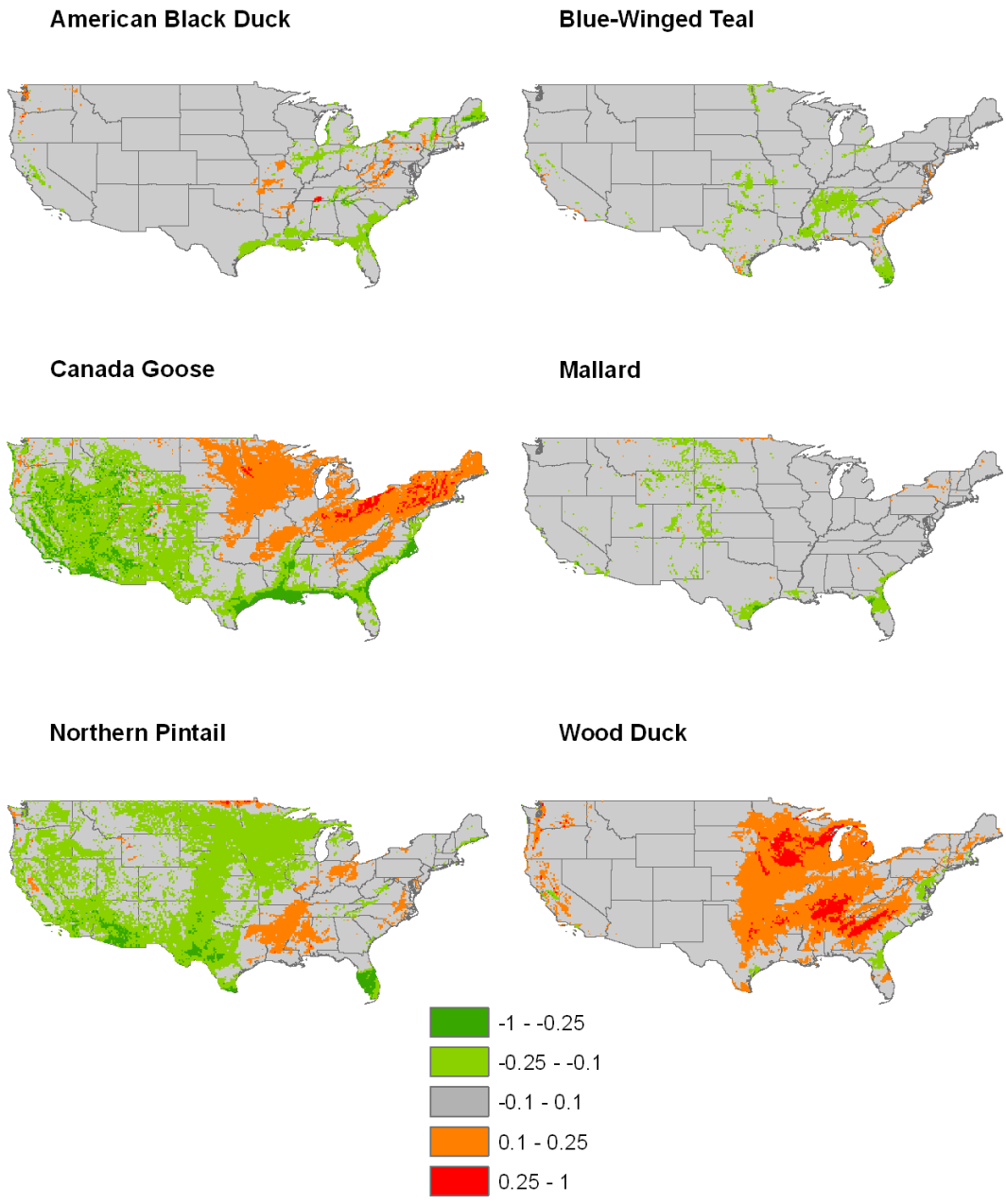

Figure B1: Maps of changes in winter 1990s predicted probability of occurrence relative to the winter 1950s predicted probability of occurrence.

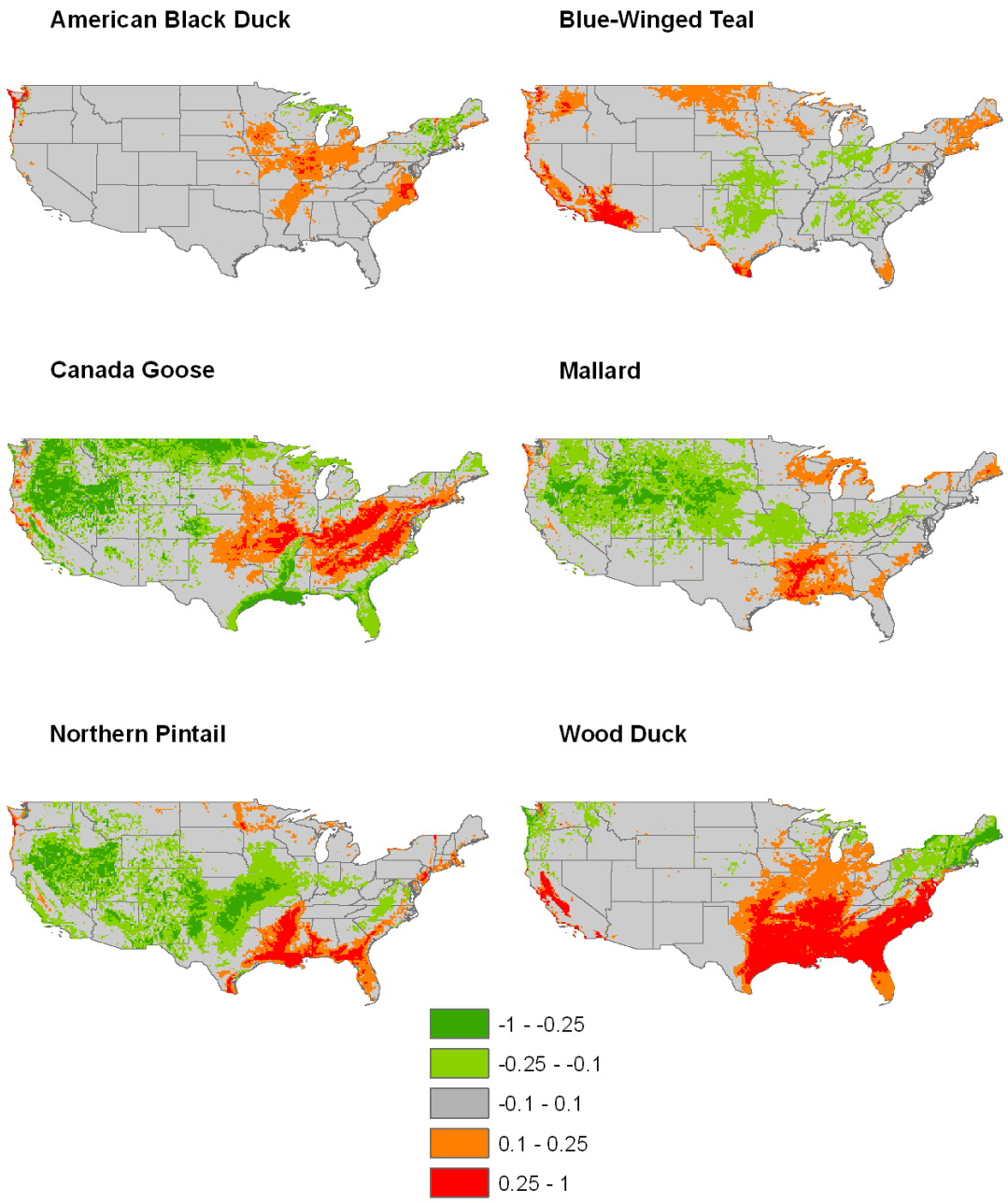

Figure B2: Maps of changes in spring 1990s predicted probability of occurrence relative to the spring 1950s predicted probability of occurrence.

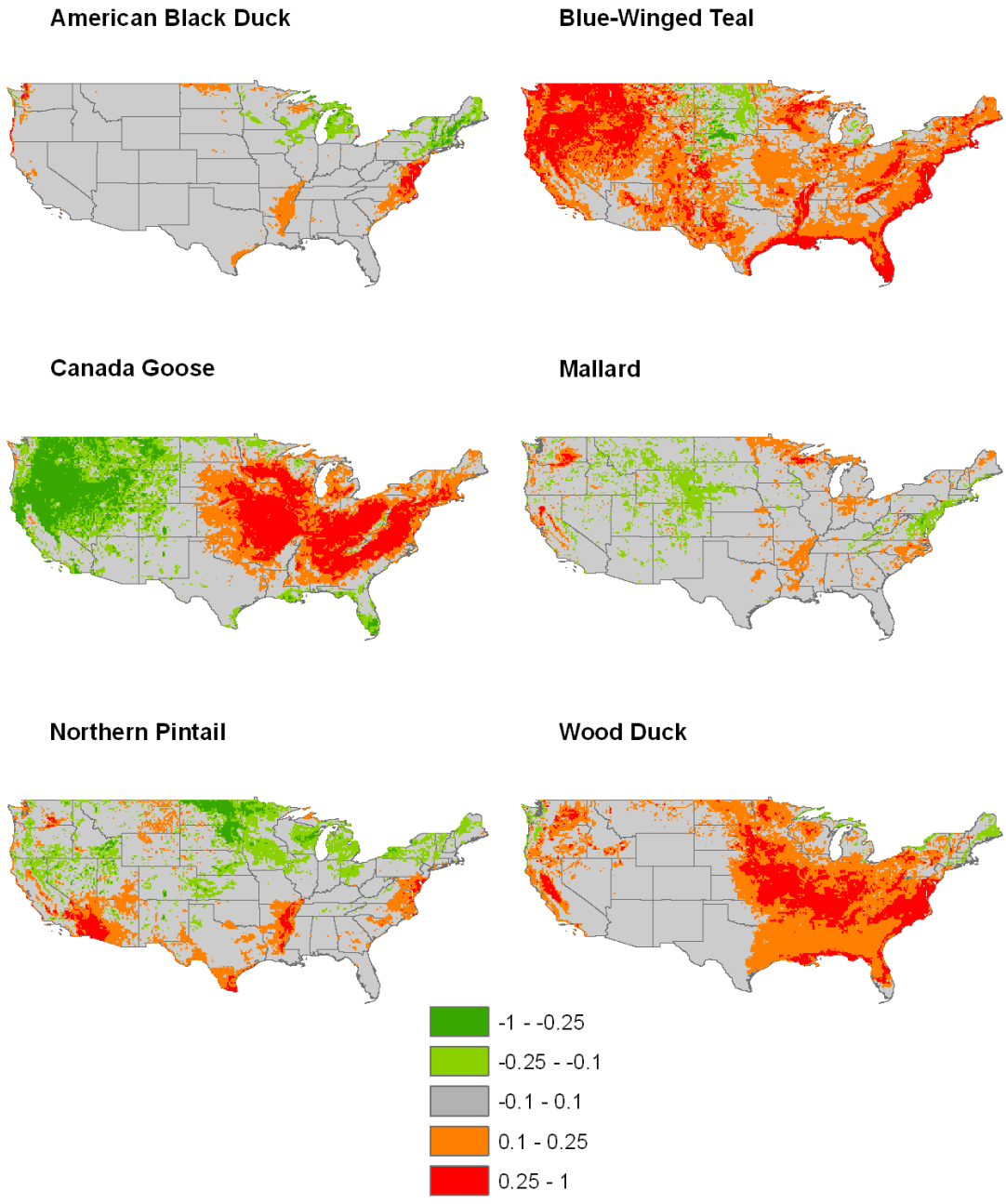

Figure B3: Maps of changes in summer1990s predicted probability of occurrence relative to the summer 1950s predicted probability of occurrence**.**

Appendix C: Chapter Three Supplemental Material Part Two

Species distribution models are frequently used to predict how species will respond to future climate changes (Midgley *et al.* 2002; Travis 2003; Lawler *et al.* 2006). Distribution models are constructed based on the relationship between known present locations of a species and the environmental conditions. The forecasted impact of climate change on species distributions are typically derived based on the assumption that the species are in equilibrium (Guisan and Zimmermann 2000). In other words, the assumption is that the species exists presently in its optimal habitat and is not adapting or really capable of adapting. To eliminate some of the uncertainty created by this assumption, a validation step is needed (Hijmans and Graham 2006; Botkin *et al.* 2007). The following work is an attempt to incorporate a validation step into the process of predicting species responses to climate change. However instead of predicting the impacts of future climate change, the 1990s distribution model of the six study species to the 1950s environment were backcasted.

For the 1950s and the 1990s, models for each study species in each season were created. This was according to the exact methods and data described in Chapter Three. The MaxEnt models used to create the 1990s probability surfaces were used to project probability surfaces according to the 1950s environmental data; again doing this for each study species and season. The 1950s projected surfaces essentially illustrate where the species would likely be present if it selected habitat according to the manner in which it did in the 1990s. The projected 1950s surfaces were compared spatially and statistically to the 1950s surfaces created from actual 1950s species occurrence data.

Table C1: Kappa statistics comparing predicted probability of occurrence from actual 1950s species data to projected probability of occurrence from the 1990s model. The corresponding standard deviation is in parentheses.

	Fall	**Winter**	**Spring**	**Summer**
American Black Duck	0.770 (0.010)	0.687 (0.015)	0.835 (0.013)	0.646 (0.018)
Blue-Winged Teal	0.406 (0.009)	0.691 (0.017)	0.624 (0.011)	0.402 (0.010)
Canada Goose	0.228 (0.010)	0.154 (0.008)	0.536 (0.009)	0.085 (0.008)
Mallard	0.478 (0.009)	0.753 (0.008)	0.600 (0.009)	0.586 (0.010)
Northern Pintail	0.745 (0.009)	0.537 (0.010)	0.469 (0.012)	0.504 (0.011)
Wood Duck	0.578 (0.011)	0.268 (0.013)	0.090 (0.010)	0.473 (0.014)

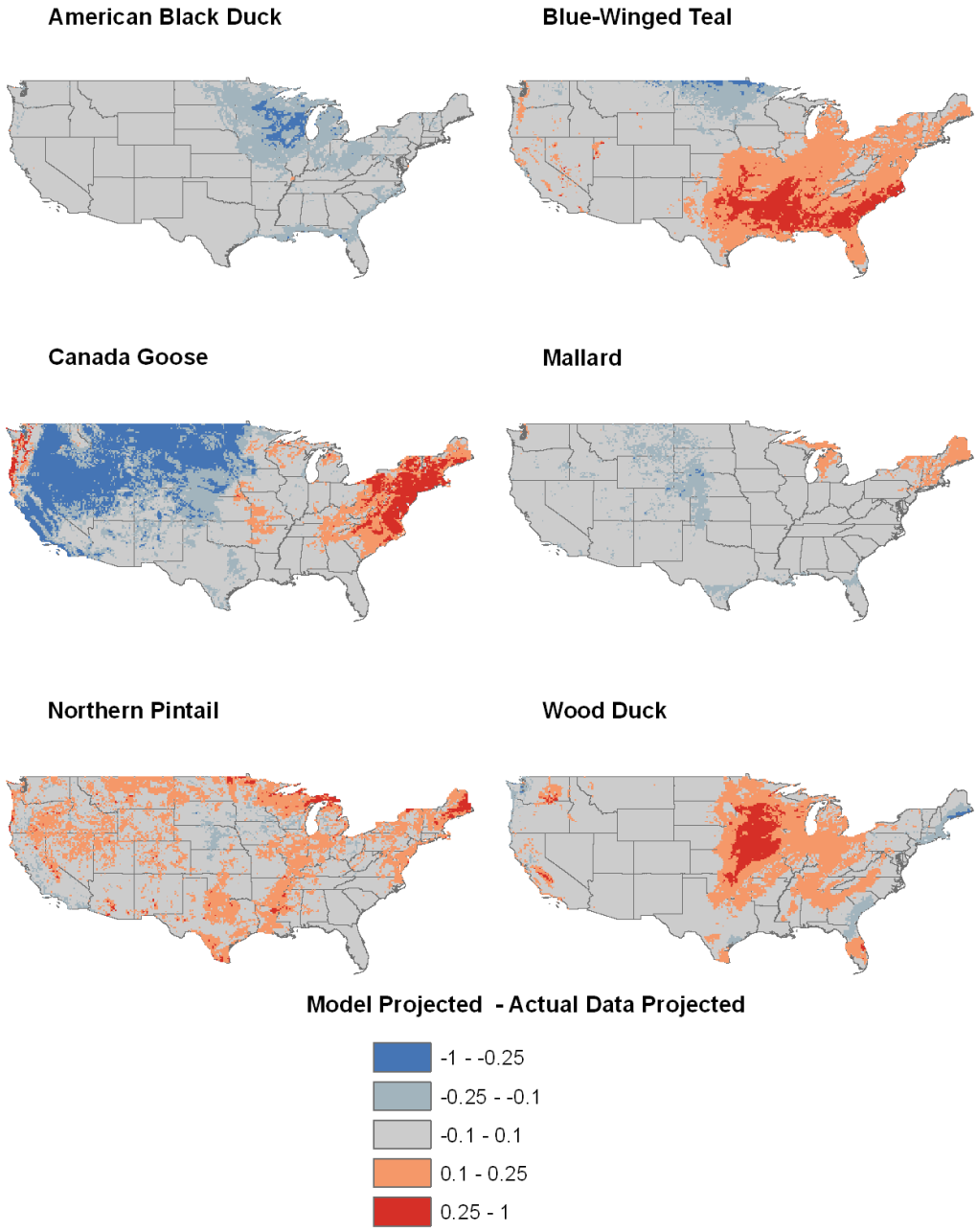

Figure C1: Map of the difference between projected results for the fall 1950s and results from actual fall 1950s data predictions.
The projected results were based on MaxEnt models created with 1990s occurrence data, and then forecasted according to 1950s environmental data. The predicted probability for the 1950s were based on MaxEnt models creates with 1950s occurrence data. The red areas illustrate areas that were over-predicted when compared to the predictions based on actual data, and the blue areas were under-predicted.

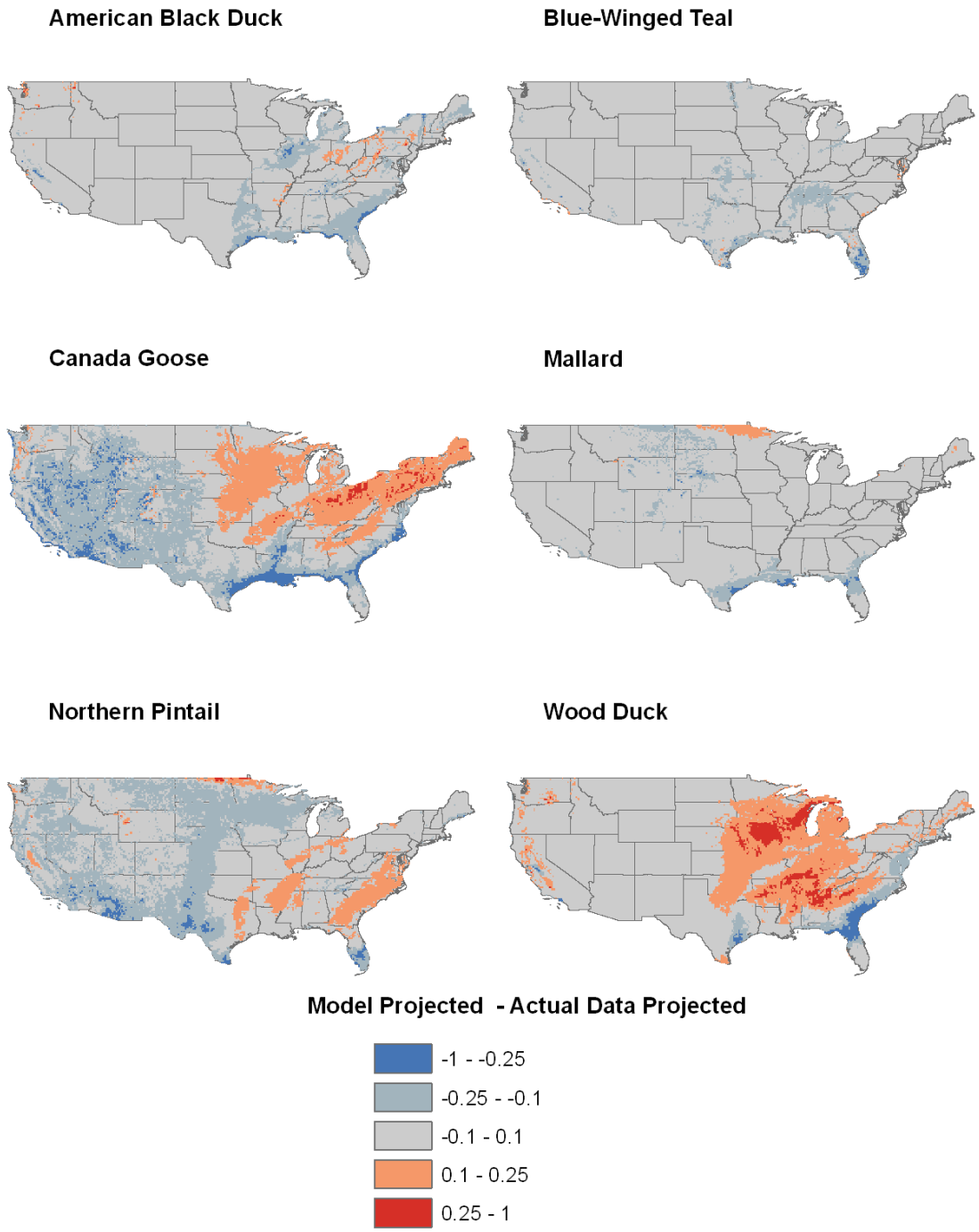

Figure C2: Map of the difference between projected results for the winter 1950s and results from actual winter 1950s data predictions.
See Figure C1 for description of methods.

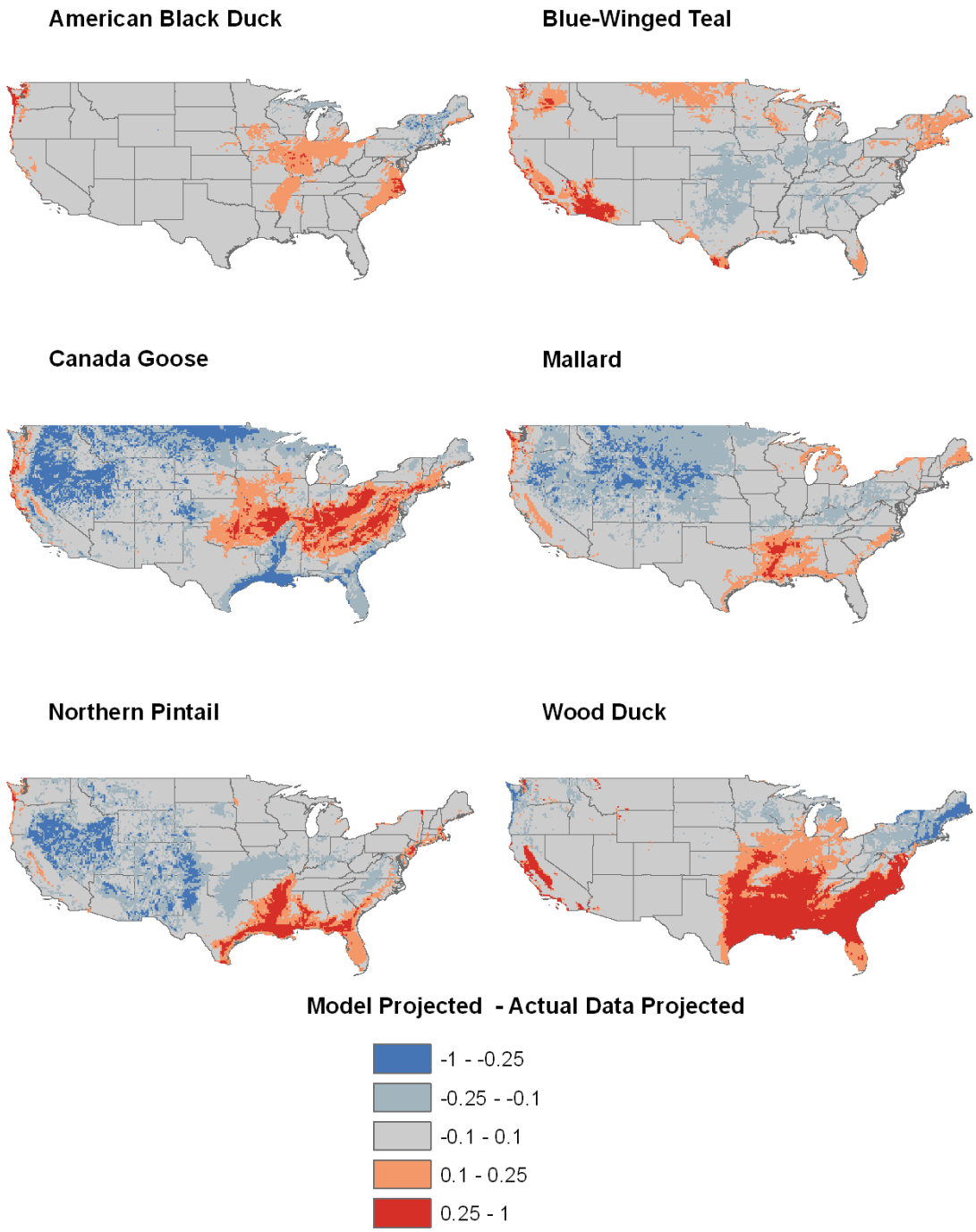

Figure C3: Map of the difference between projected results for the spring 1950s and results from actual spring 1950s data predictions**.**
See Figure C1 for description of methods.

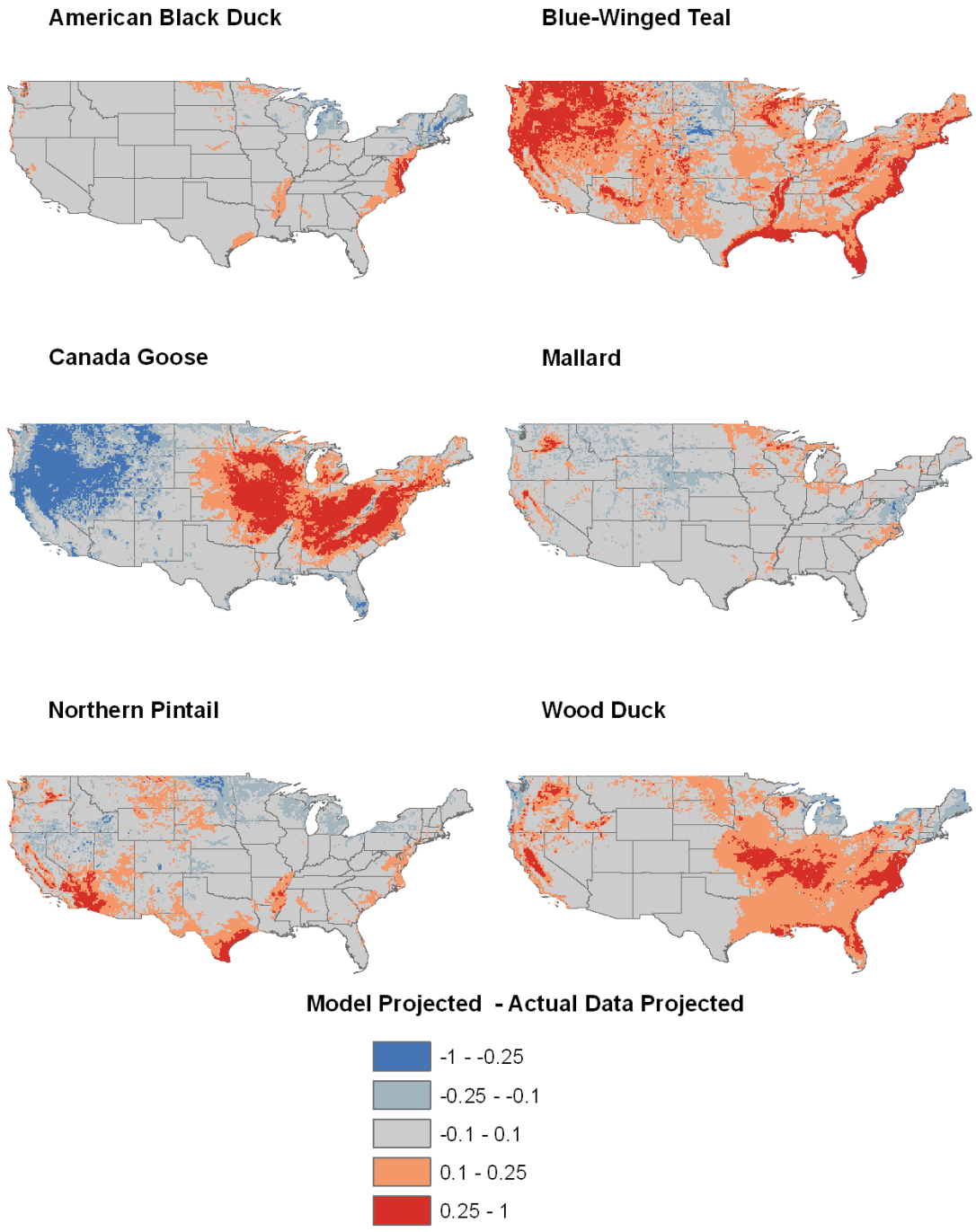

Figure C4: Map of the difference between projected results for the summer 1950s and results from actual summer 1950s data predictions.
See Figure C1 for description of methods.

Appendix D: Chapter Four Supplemental Material

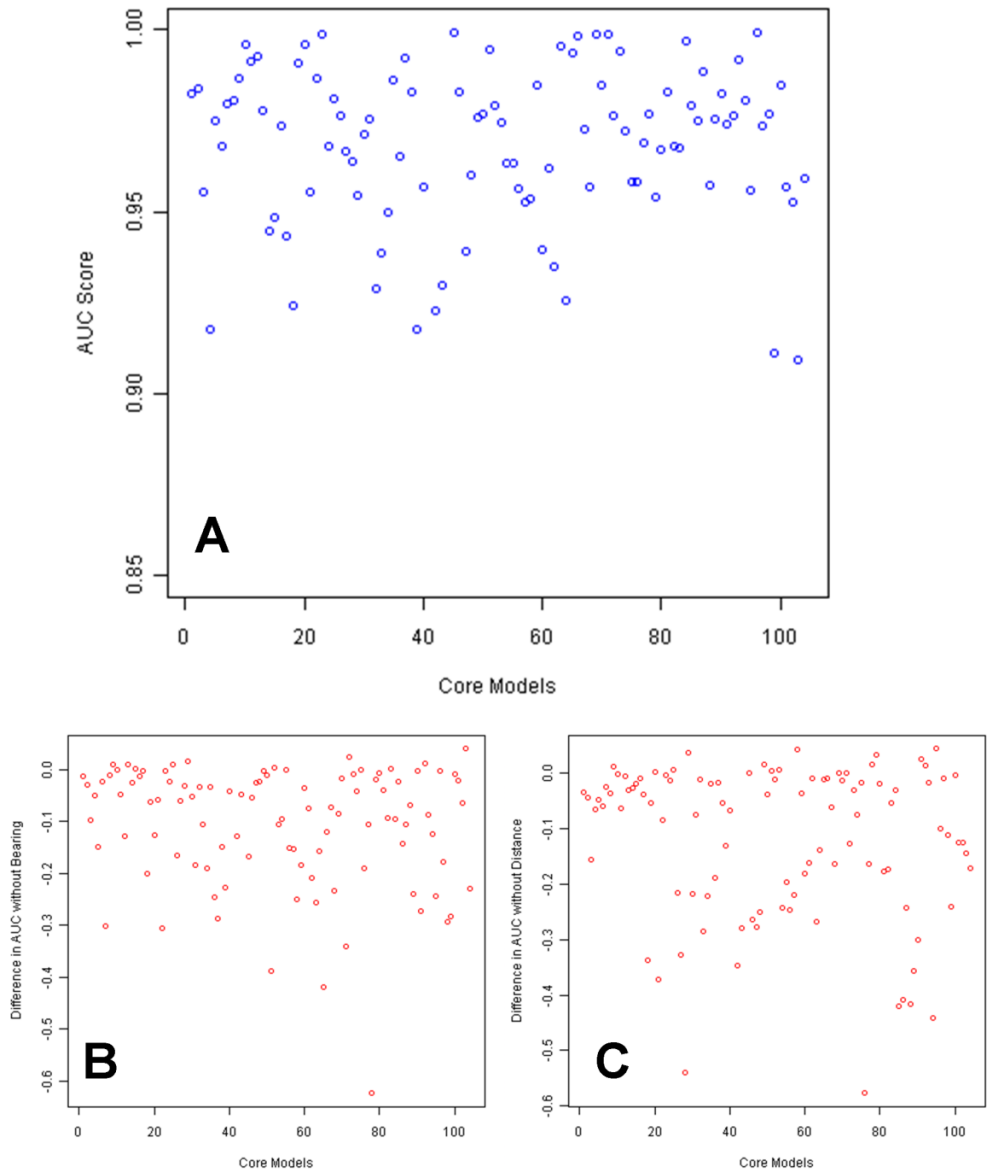

Figure D1: Scatterplots of core model AUC scores.
Figure D1-A: AUC scores for each of the 104 core models for the fall blue-winged teal case study. Figure D1-B: The difference in AUC scores for core models built without the bearing variable compared to AUC scores of the full model. Figure D1-C: The difference in AUC scores for core models built without the distance variable compared to AUC scores of the full model.

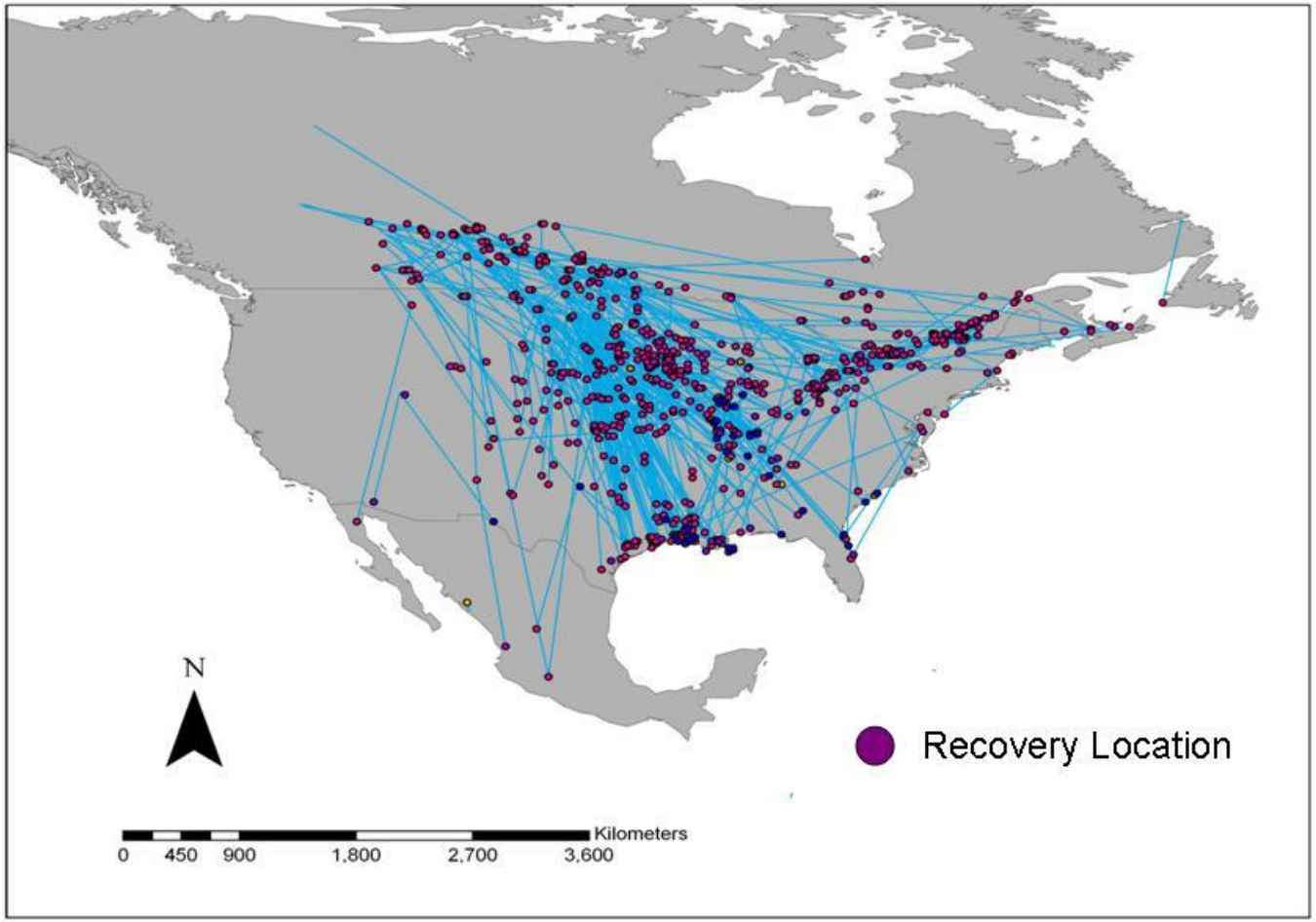

Figure D2: Map of within-10-day recovered blue-winged teal.
The purple points illustrate the recovery locations of blue-winged teal individuals that were recovered within 10 days of being banded. The blue lines illustrate the shortest Euclidean distance path between the banding and recovery location. The intent of this figure is (1) to demonstrate that there are large amounts of this quickly recovered data, and (2) to show how movement information can be gathered.

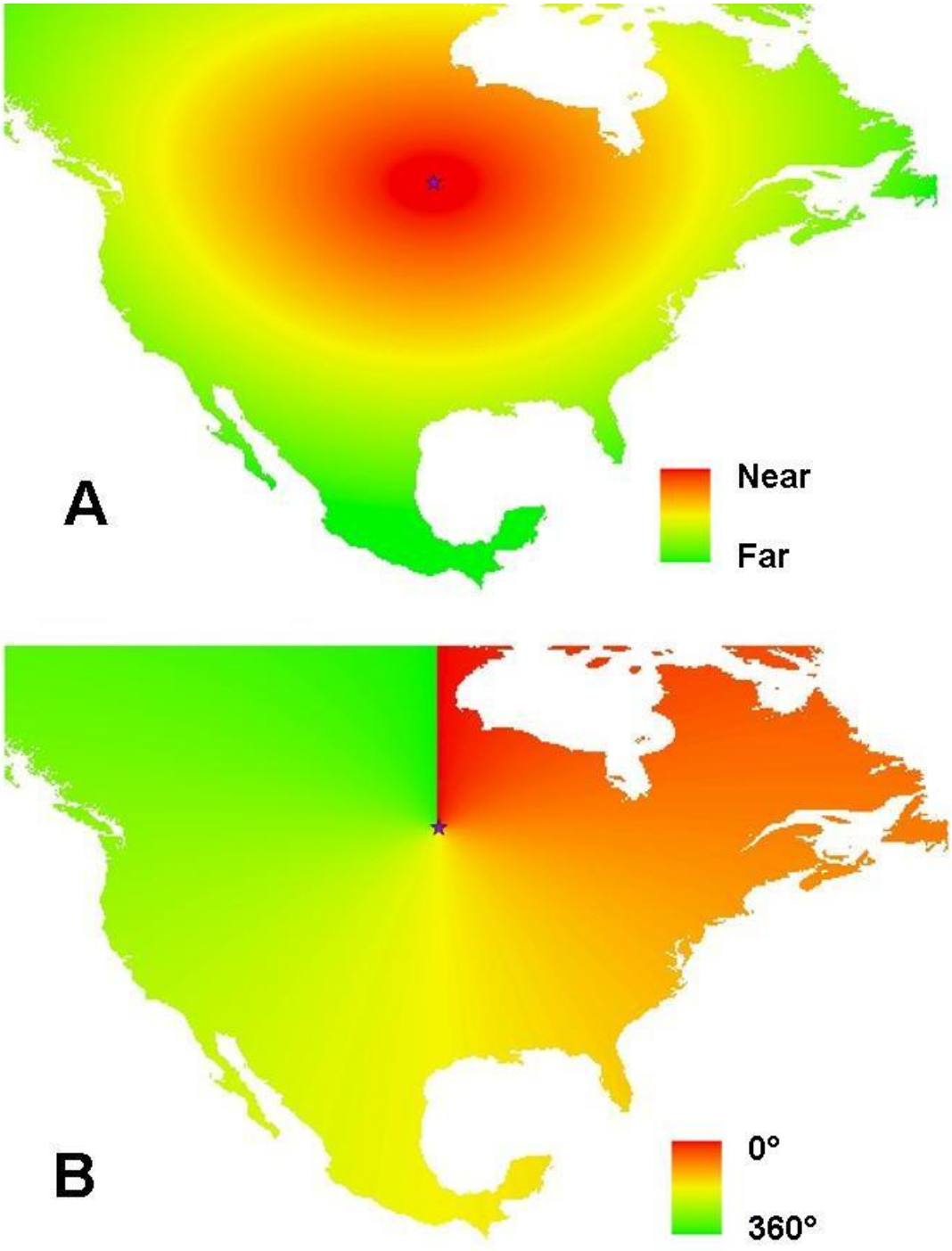

Figure D3: Map illustrating the measurement of distance and bearing from a point.
Figure D3-A: Map illustrating how great circle distance is measured across the study area from a given point (here depicted with a purple star). The measurements for distance would be the great circle distance a within-30-recoveried point is from a location of a core model. Figure D3-B: Map illustrating how bearing is measured across the study area from the same point in Figure D3-A.

Glossary

10-Minute Grid: Spatial grid that the USGS Patuxent Wildlife Research Center Bird Banding Laboratory uses to report banding and recovery locations. This grid is approximately a 16km resolution.

American Black Duck (*Anas rubripes*): Large dabbling duck of Northeast America (The Cornell Lab of Ornithology 2011). Found almost exclusively along the east coast of the United States, but individuals have been found as far as France. American black duck is one of the six study species selected for this research.

Area under the Curve (AUC): Measure of the area under a receiver operating characteristic (ROC) curve. This measure provides insight into the performance of a model. An AUC score of 0.5 would be equivalent to random, while an AUC of 1.0 would be a model that is a perfect predictor. For this research, an AUC of 0.7 was the threshold for a model to be considered acceptable.

Banding (or Band): Refers to the point in time where the individual bird was captured and given its identifying band. Banding would also be synonymous with marking in reference to mark-recapture studies. For this research specifically, only banding events recorded by the North American Bird Banding Program were used.

Bird Banding Laboratory (BBL): Part of the North American Bird Banding Program along with Canada Wildlife Service. The BBL is within the USGS Patuxent Wildlife Research Center, and is responsible for maintaining banding-recovery records as well as organizing the work of approximately 6,100 banders in the United States and Canada.

Blue-Winged Teal (*Anas discors*): One of the smaller species of dabbling ducks. Blue-winged teal is found year-round throughout North America and in South America during the winter season. Blue-winged teal is one of the six study species selected for this research.

Canada Goose (*Branta canadensis*): The largest of the study species, and the only goose species selected for this research. Canada goose can be found throughout the United States and Canada and are frequently considered a nuisance species.

Core Models: Used to refer to the movement models in chapter four that were created with multiple within-30-day recoveries. These core models are then used to project the movement from points that lacked adequate recoveries.

Depth to Water Table (DWT): Wetland proxy data set, which is a simulated long-term measure of the point either at or below the surface where climate and geological/topographic water fluxes balance.

Generalized Linear Model (GLM): Used to relate species occurrence data to independent environmental data. GLMs are based on linear regressions created by using a link function. Specifically for this research, GLMs were fit using a binomial distribution and logistic link.

Great Circle Distance: Measure of the distance between any two points along the surface of a sphere.

Kappa Statistic: Measures the agreement between two data sets. For example, this research used kappa statistics to measure the agreement between model projections for bird distributions of the 1950s and 1990s. While taking chance agreement into account, the kappa statistic encapsulates all the information supplied in a confusion matrix of a classification. A kappa statistic of 1.0 would be perfect agreement between the two data sets.

Mallard (*Anas platyrhynchos*): Large dabbling duck that can be found throughout North America and Eurasia. Mallard is beginning to out compete several other species, including blue-winged teal. Mallard is one of the six study species selected for this research.

Maximum Entropy (MaxEnt): Method of species distribution modeling which attempts to find the target distribution that is closest to uniform and constrained only by the empirical averages of the independent data. MaxEnt approach to distribution modeling does not require absences or pseudo-absences; therefore it is well-suited for presence-only data like the BBL database.

Nonmetric Multidimensional Scaling (NMS): An ordination technique that iteratively fits the data into a best fit ordination space, which is constructed from the multivariate environmental data. NMS uses a step-down procedure to find the most appropriate ordination dimensionality. The data points are systematically moved around the ordination space until a minimum stress value, a measure of distance between the points in environmental and ordination space, is achieved.

Northern Pintail (*Anas acuta*): Medium sized duck that feeds on insects, grain, and other plants. Northern pintail is a common species in North and Central America. Northern pintail is one of the six study species selected for this research.

Phenology: The study of the timing of natural events

Predicted Probability of Occurrence: The chance the species is likely to occur in any given cell in the study area based on the distribution model. In this case, all predictions are based on models that were created from actual independent data.

Projected Probability of Occurrence: The chance the species is likely to occur in any given cell in the study area based on a distribution model. In this case, the projected predictions are based on a model created with different independent data. The assumption is that the probability distribution of species occurrence is unchanged to the new environmental conditions.

Pseudo-Absence: The selection of background points to serve as assumed absences for a species. The creation of pseudo-absences is necessary when a distribution modeling method requires information about the non-habitat area and known absences of the species are not available.

Random Forest: Method of species distribution modeling. This method creates multiple classification trees of habitat and non-habitat. Each tree is grown with a certain percentage of the data, and OOB data are used to measure the accuracy.

Receiver Operating Characteristic Curve (ROC): Plot of the true positive rate (sensitivity) against the false positive rate (1-specificity). This plot is used to calculate the AUC score of a model's performance.

Recovery: Refers to every time a banded bird is recorded after the initial banding. For the BBL data, often, but not always, this is usually a recovery event by a hunter. Recovery events are synonymous with recaptures.

Species Distribution Modeling: Methods used to predict the entire spatial distribution of species. These techniques are often employed since it is often impossible to have complete knowledge of a species location. These techniques are also used to make predictions about how a species may respond to changes in the environment.

Species Specific Annual Cycle: Timing of annual events (i.e. fall migration, winter/non-breeding, spring migration, summer/breeding) specifically for each of the study species.

Wood Duck (*Aix sponsa*): Cavity-nesting dabbling duck. Wood duck can be found year-round in the southeastern portion of the United States. This is the only study species with bottomland forests/swamps as habitat.

Within-30-Day Recovery: Recovery of banded birds that fall within 30 days of the banding event. These recovery events were used to determine periods of the year

when the study species had the greatest directed velocity; essentially to determine migration periods.

Table E1: List of commonly used abbreviations

Study Species		
abd	American black duck	*Anas rubripes*
bwt	blue-winged teal	*Anas discors*
cg	Canada goose	*Branta canadensis*
mal	Mallard	*Anas platyrhynchos*
np	northern pintail	*Anas acuta*
wd	wood duck	*Aix sponsa*

AUC	Area under the Curve
BBL	Bird Banding Laboratory
DWT	Depth to Water Table
Elev	Elevation
GLM	Generalized Linear Model
GPS	Global Positioning System
MaxEnt	Maximum Entropy
NLCD	National Land Cover Database
NMS	Nonmetric Multidimensional Scaling
OOB	Out-Of-Bag
PPR	Prairie Pothole Region
PPT	Precipitation
PW	Percent Wetland
ROC	Receiver Operating Characteristic
SRTM	Shuttle Radar Topography Mission
Temp	Temperature
USA	United States of America
USGS	United States Geological Survey

References

Acevedo, P., M.Á. Farfán, A.L. Márquez, M. Delibes-Mateos, R. Real, and J.M.Vargas. 2010. Past, present and future of wild ungulates in relation to changes in land use. *Landscape Ecology* 26: 19-31.

Ackerman, J,T., J.Y. Takekawa, D.L. Orthmeyer, J.P. Fleskes, J.L. Yee, and K.L.Kruse. 2006. Spatial use by wintering greater white-fronted geese relative to a decade of habitat change in California's Central Valley. *Journal of Wildlife Management* 70: 965-976.

Adrianensen, F., J.P. Chardon, G. De Blust, E. Swinnen, S. Villalba, H. Gulinck, and E. Matthysen. 2003. The application of 'least-cost' modelling as a functional landscape model. *Landscape and Urban Planning* 64: 233-247.

Alerstam, T. 1996. The geographical scale factor in orientation of migrating birds. *Journal of Experimental Biology* 199: 9 -19.

———. 2001. Detours in Bird Migration. *Journal of Theoretical Biology* 209: 319-331.

Alerstam, T, G.A. Gundmundsson, P.E. Jonsson, J. Karlsson, and A. Lindstrom. 1990. Orientation, Migration Routes and Flight Behaviour of Knots, Turnstones and Brant Geese Departing from Iceland in Spring. *Artic* 43: 201-214.

Alerstam, T. and A. Hedenström. 1998. The Development of Bird Migration Theory. *Journal of Avian Biology* 29: 343-369.

Andrewarthe, H.G., and L.C. Birch. 1954. *The Distribution and Abundance of Animals*. Chicago, IL: University of Chicago Press.

Araújo, M.B., and R.G. Pearson. 2005. Equilibrium of species' distributions with climate. *Ecography* 28: 693-695.

Audubon, M.R., J.J. Audubon, and E. Coues. 1994. *Audubon and His Journals*. Courier Dover Publications,

Bairlein, F. 2003. The study of bird migrations – some future perspectives: Capsule Routes and destinations have been unveiled but modern techniques offer the chance to explore much more. *Bird Study*: 243.

———. 2008. The mysteries of bird migration-still much to be learnt. *British Birds* 101: 68-81.

Bartel, R.A, and J.O. Sexton. 2009. Monitoring habitat dynamics for rare and endangered species using satellite images and niche-based models. *Ecography* 32: 888-896.

Bartzen, B.A., K.W. Dufour, R.G. Clark, and F.D. Caswell. 2011. Trends in agricultural impact and recovery of wetlands in prairie Canada. *Ecological Applications* 20: 525-538.

Bell, J.F., and A.H. Fielding. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation* 24: 38-49.

Bellrose, F. 1976. *Ducks, Geese, and Swans of North America. Second Edition.* Harrisburg, PA: Stackpole Books.

Bent, A.C. 1987. *Life Histories of North American Wild Fowl*. New York: Dover Publications.

Berthold, P., and S.B. Terrill. 1991. Recent Advances in Studies of Bird Migration. *Annual Review of Ecology and Systematics* 22: 357-378.

Bleeker, C. J. 1964. The Pattern of the Ancient Egyptian Culture. *Numen* 11: 75-82.

Bolduc, F., and A.D. Afton. 2008. Monitoring waterbird abundance in wetlands: The importance of controlling results for variation in water depth. *Ecological Modelling* 216: 402-408.

Botkin, D.B., H. Saxe, M.B. Araujo, R. Betts, R.H.W. Bradshaw, T. Cedhagen, P. Chesson, et al. 2007. Forecasting the effects of global warming on biodiversity." *Biosciences* 57: 227-236.

Breiman, L. 1984. *Classification and regression trees*. Wadsworth International Group.

Breiman, L. 2001. Random Forests. *Machine Learning* 45 (1): 5-32.

Brotons, L., W. Thuiller, M.B. Araújo, and A.H. Hirzel. 2004a. Presence-absence versus presence-only modelling methods for predicting bird habitat suitability. *Ecography* 27: 437-448.

Buckley, P.A., C.M. Francis, P. Blancher, D.F. DeSante, C.S. Robbins, G. Smith, and P. Cannell. 1998. The North American Bird Banding Program: Into the 21st century. *Journal of Field Ornithology* 69: 511-529.

Cameron, S.E., J.L. Parra, P.G. Jones, A. Jarvis, and R.J. Hijmans. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25: 1965-1978.

Canepuccia A.D., J.P. Isacch, D.A. Gagliardini, A.H. Escalante, and O.O. Iribarne. 2007. Waterbird Response to Changes in Habitat Area and Diversity Generated by Rainfall in a SW Atlantic Coastal Lagoon. Waterbirds 30:541-553.

Casas, F., F. Mougeot, J. Vinuela, and V. Bretagnolle. 2009. Effects of hunting on the behaviour and spatial distribution of farmland birds: importance of hunting-free refuges in agricultural areas. *Animal COnservation* 12: 346-354.

Clarke, K.R. 1993. Non-parametric multivariate analyses of changes in community structure. *Australian Journal of Ecology* 18: 117-143.

Conover, M. R., and G.G. Chasko. 1985. Nuisance Canada Goose Problems in the Eastern United States. *Wildlife Society Bulletin* 13: 228-233.

Coops, N.C., M.A. Wulder, and D. Iwanicka. 2009. Exploring the relative importance of satellite-derived descriptors of production, topography and land cover for predicting breeding bird species richness over Ontario, Canada. *Remote Sensing of Environment* 113: 668-679.

Crick, H.Q.P. 2004. The impact of climate change on birds. *Ibis* 146: 48-56.

Cutler, D.R., T.C. Edwards, K.H. Beard, A. Cutler, K. T. Hess, J. Gibson, and J.J. Lawler. 2007. Random Forests for Classification in Ecology. *Ecology* 88: 2783-2792.

Dahl, T.E. 1990. Wetland losses in the United States 1780's to 1980's. U.S. Department of the Interior Fish and Wildlife Service, Washington, DC, USA.

Dahl, T.E. 2006. Status and trends of wetlands in the conterminous United States 1998 to 2004. US Department of Interior, Fish and Wildlife Service, Washington, DC. USA, p. 112.

Desgranges, J., J. Ingram, B. Drolet, J. Morin, C. Savage, and D. Borcard. 2006. Modelling Wetland Bird Response to Water Level Changes in the Lake Ontario – St. Lawrence River Hydrosystem. *Environmental Monitoring and Assessment* 113: 329-365.

Desholm, M. 2003. How much do small-scale changes in flight direction increase overall migration distance? *Journal of Avian Biology* 34: 155-158.

Dixon, P. 2003. Vegan, A Package of R Functions for Community Ecology. *Journal of Vegetation Science* 14: 927-930.

Downs, J.A., and M.W. Horner. 2008. Spatially modelling pathways of migratory birds for nature reserve site selection. *International Journal of Geographical Information Science* 22: 687.

Driezen, K., F. Adriaensen, C. Rondinini, C.P. Doncaster, and E. Matthysen. 2007. Evaluating least-cost model predictions with empirical dispersal data: A case-study using radiotracking data of hedgehogs (*Erinaceus europaeus*). *Ecological Modelling* 209: 314-322.

Dudei, N.L., and A.L. Stigall. 2010. Using ecological niche modeling to assess biogeographic and niche response of brachiopod species to the Richmondian Invasion (Late Ordovician) in the Cincinnati Arch. *Palaeogeography, Palaeoclimatology, Palaeoecology*: 28-43.

Dudík, M., S.J. Phillips, and R.E. Schapire. 2007. Maximum Entropy Density Estimation with Generalized Regularization and an Application to Species Distribution Modeling. *The Journal of Machine Learning Research* 8: 1217–1260.

Earnst, S.L. 1994. Tundra Swan Habitat Preferences during Migration in North-Dakota. *Journal of Wildlife Management* 58: 546-551.

Elith, J., C.H. Graham, R.P. Anderson, M. Dudík, S. Ferrier, A. Guisan, R. J. Hijmans, et al. 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29: 129-151.

Elith, J., S.J. Phillips, T. Hastie, M. Dudík, Y.E. Chee, and C.J. Yates. 2011. A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions* 17: 43-57.

Erwin, R.M. 2002. Integrated Management of Waterbirds: Beyond the Conventional. *Waterbirds: The International Journal of Waterbird Biology* 25: 5-12.

Esler, D., D.M. Mulcahy, and R.L. Jarvis. 2000. Testing Assumptions for Unbiased Estimation of Survival of Radiomarked Harlequin Ducks. *The Journal of Wildlife Management* 64: 591-598.

Faaborg, J., R.T. Holmes, A.D. Anders, K.L. Bildstein, K.M. Dugger, S.A. Gauthreaux, P. Heglund, et al. 2010. Recent advances in understanding migration systems of New World land birds. *Ecological Monographs* 80: 3-48.

Fan, Y., and G. Miguez-Macho. 2010. A simple hydrologic framework for simulating wetlands in climate and earth system models. *Climate Dynamics*.

Fan, Y., G. Miguez-Macho, C.P. Weaver, R. Walko, and A. Robock. 2007. Incorporating water table dynamics in climate modeling: 1. Water table observations and equilibrium water table simulations. *Journal of Geophysical Research* 112: 17.

Farmer, A.H., and A.H. Parent. 1997. Effects of the Landscape on Shorebird Movements at Spring Migration Stopovers. *The Condor* 99: 698-707.

Fox, A.D., C.M. Glahder, and A.J. Walsh. 2003. Spring migration routes and timing of Greenland white-fronted geese - results from satellite telemetry. *Oikos* 103: 415-425.

Franklin, J. 2010. Moving beyond static species distribution models in support of conservation biogeography. *Diversity and Distributions* 16: 321-330.

Freeman, E. 2007. PresenceAbsence: An R Package for Presence-Absence Model Evaluation. *USDA Forest Service, Rocky Mountain Research Station*.

Fretwell, J.D., J.S. Williams, P.J. Redman, and Geological Survey (U.S.). 1996. *National water summary on wetland resources*. U.S. G.P.O.

Fujita, G., G. Hong-Liang, M. Ueta, O. Goroshko, V. Krever, K. Ozaki, N. Mita, and H. Higuchi. 2004. Comparing areas of suitable habitats along travelled and possible shortest routes in migration of White-naped Cranes Grus vipio in East Asia. *Ibis* 146: 461-474.

Gauthreaux, S.A. 1979. Priorities in Bird Migration Studies. *The Auk* 96: 813-815.

Gilbert, M., P. Chaitaweesub, T. Parakamawongsa, S. Premashthira, T. Tiensin, W. Kalpravidh, H. Wagner, and J. Slingenbergh. 2006. Free-grazing Ducks and Highly Pathogenic Avian Influenza, Thailand. *Emerging Infectious Diseases* 12: 227-234.

Gill, Frank B. 1995. *Ornithology*. Macmillan.

Gilmer, D. S., I.J. Ball, L.M. Cowardin, and J. H. Riechmann. 1974. Effects of Radio Packages on Wild Ducks. *The Journal of Wildlife Management* 38: 243-252.

Gómez-Rodríguez, C., J. Bustamante, S. Koponen, and C. Díaz-Paniagua. 2008. High-resolution remote-sensing data in amphibian studies: identification of breeding sites and contribution to habitat models. *The Herpetological Journal* 18 (April): 103-113.

Gonzalez, M.L, and J.W Hollister. 2008. Assessing the Accuracy of National Land Cover Dataset Area Estimates at Multiple Spatial Extents. Photogrammetric engineering and remote sensing. *Photogrammetric engineering and remote sensing* 70: 405-414.

Graham, C.H, J. VanDerWal, S.J. Phillips, C. Moritz, and S.E Williams. 2010. Dynamic refugia and species persistence: tracking spatial shifts in habitat through time. *Ecography* 33: 1062-1069.

Green, A.J., J. Figuerola, and M.I. Sanchez. 2002. Implications of waterbird ecology for the dispersal of aquatic organisms. *ACTA Oecologica-Inernational Journal of Ecology* 23: 177-189.

Greenwood, J.D. 2009. 100 years of ringing in Britain and Ireland. *Ringing & Migration*: 147.

Grier, J., and D. Johnson. 1988. Determinants of Breeding Distribution of Ducks. *USGS Northern Prairie Wildlife Research Center* (January 1).

Grinnell, J.. 1917. Field Tests of Theories Concerning Distributional Control. *The American Naturalist* 51: 115-128.

Guisan, A., and W. Thuiller. 2005. Predicting species distribution: offering more than simple habitat models. *Ecology Letters* 8: 993-1009.

Guisan, A., and N.E. Zimmermann. 2000. Predictive habitat distribution models in ecology. *Ecological Modelling* 135: 147-186.

Guisan, A, T.C. Edwards, and T. Hastie. 2002. Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecological Modelling* 157: 89-100.

Gustafson, M.E., and M E. Hildenbrand. 1999. Bird Banding Laboratory homepage, version 07–10–2005 *(*http://www.pwrc.usgs.gov/bbl/*)*.

Haig, S.M., D.W. Mehlman, and L.W. Oring. 1998. Avian movements and wetland connectivity in landscape conservation. *Conservation Biology* 12: 749-758.

Hedenström, A., and T. Alerstam. 1997. Optimum Fuel Loads in Migratory Birds: Distinguishing Between Time and Energy Minimization. *Journal of Theoretical Biology* 189: 227-234.

Heikkinen, R.K., M. Luoto, M.B. Araújo, R. Virkkala, W. Thuiller, and M.T. Sykes. 2006. Methods and uncertainties in bioclimatic envelope modelling under climate change. *Progress in Physical Geography* 3: 751 -777.

Hijmans, R. J., E Williams, and C Vennes. 2011. geosphere: R package.

Hijmans, R.J., and C.H. Graham. 2006. The ability of climate envelope models to predict the effect of climate change on species distributions. *Global Change Biology* 12: 2272-2281.

Hitch, A.T., and P.L. Leberg. 2007. Breeding Distributions of North American Bird Species Moving North as a Result of Climate Change. *Conservation Biology* 21: 534-539.

Holt, R.D. 1992. A neglected facet of island biogeography: The role of internal spatial dynamics in area effects. *Theoretical Population Biology* 41: 354-371.

Homer, C., C. Huang, L. Yang, B. Wylie, and M. Coan. 2004. Development of a 2001 National Landcover Database for the United States. *Photogrammetric Engineering and Remote Sensing* 70: 829-840.

Hulme, P.E. 2005. Adapting to climate change: is there scope for ecological management in the face of a global threat? *Journal of Applied Ecology* 42: 784-794.

Hunerbein, K. von, H.J. Hamann, E. Ruter, and W. Wiltschko. 2000. A GPS-based system for recording the flight paths of birds. *Naturwissenschaften* 87: 278-279.

Hutto, R. L. 1998. On the Importance of Stopover Sites to Migrating Birds. *The Auk* 115: 823-825.

———. 2000. On the importance of en route periods to the conservation of migratory landbirds. *The Auk* 115: 823-825.

Iverson, L.R., and A.M. Prasad. 1998. Predicting Abundance of 80 Tree Species Following Climate Change in the Eastern United States. *Ecological Monographs* 68: 465-485.

Johnson, W.C., B.V. Millett, T. Gilmanov, R.A. Voldseth, G.R. Guntenspergen, and D.E. Naugle. 2005. Vulnerability of Northern Prairie Wetlands to Climate Change. *BioScience* 55: 863.

Johnson, W.C., B. Werner, G.R. Guntenspergen, R.A. Voldseth, B. Millett, D.E. Naugle, M. Tulbure, R.W.H. Carroll, J. Tracy, and C. Olawsky. 2010. Prairie Wetland Complexes as Landscape Functional Units in a Changing Climate. *BioScience* 60: 128-140.

Kaufman, K. 1996. *Lives of North American Birds*. New York: Houghton Mifflin Company.

Kölzsch, A., and B. Blasius. 2008. Theoretical approaches to bird migration. *The European Physical Journal - Special Topics* 157: 191-208.

Koneff, M.D., and J.A. Royle. 2004. Modeling wetland change along the United States Atlantic Coast. *Ecological Modelling* 177: 41-59.

Krapu, G., A. Klett, and D. Jorde. 1983. The Effect of Variable Spring Water Conditions on Mallard Reproduction. DigitalCommons@University of Nebraska - Lincoln, July 1. http://digitalcommons.unl.edu/usgsnpwrc/47.

Kreakie, B.J., Y. Fan, and T. Keitt. *in review*. Enhanced Migratory Waterfowl Distribution Modeling by Inclusion of Depth to Water Table Data. *Plos One*

Landis, J.R., and G.G. Koch. 1977. The measurement of observer agreement for categorical data. *Biometrics* 33: 159-74.

Lawler, J.J, D. White, R.P. Neilson, and A.R. Blaustein. 2006. Predicting climate-induced range shifts: model differences and model reliability. *Global Change Biology* 12: 1568-1584.

Lemoine, N., H. Schaefer, and K. Böhning-Gaese. 2007. Species richness of migratory birds is influenced by global climate change. *Global Ecology and Biogeography* 16: 55-64.

De Leon, M.T., and L.M. Smith. 1999. Behavior of migrating shorebirds at North Dakota Prairie Potholes. *The Condor* 101: 645-654.

Leu, M., and C.W. Thompson. 2002. The potential importance of migratory stopover sites as flight feather molt staging areas: a review for neotropical migrants. *Biological Conservation* 106: 45-56.

Liaw, A, and M Wiener. 2002. "Classification and Regression by randomForest." *R News* 2 (3): 18-22.

Lincoln, F.C. 1928. A Bibliography of Bird Banding in America. *The Auk* 45 (January 1): 1-73.

Livezey, B.C. 1996. A Phylogenetic Analysis of Geese and Swans (Anseriformes: Anserinae), Including Selected Fossil Species. *Systematic Biology* 45: 415 -450.

———. 1997. A phylogenetic analysis of basal Anseriformes, the fossil Presbyornis, and the interordinal relationships of waterfowl. *Zoological Journal of the Linnean Society* 121: 361-428.

Loo, S.E., R. Mac Nally, and P.S. Lake. 2007. Forecasting New Zealand Mudsnail Invasion Range: Model Comparisons Using Native and Invaded Ranges. *Ecological Applications* 17: 181-189.

MacArthur, RM. 1972. *Geographical Ecology*. New York: Harper and Row.

Manel, S., H.C. Williams, and S.J. Ormerod. 2001. Evaluating presence–absence models in ecology: the need to account for prevalence. *Journal of Applied Ecology* 38: 921-931.

Marion, W.R., and J.D. Shamis. 1977. An Annotated Bibliography of Bird Marking Techniques. *Bird-Banding* 48: 42-61.

McPherson, J.M., and W. Jetz. 2007. Effects of species' ecology on the accuracy of distribution models. *Ecography* 30: 135-151.

McWilliams, S.R, C. Guglielmo, B. Pierce, and M. Klaassen. 2004. Flying, fasting, and feeding in birds during migration: a nutritional and physiological ecology perspective. *Journal of Avian Biology* 35: 377-393.

Mech, L. David. 1983. *Handbook of animal radio-tracking*. U of Minnesota Press.

Meyburg, B.U., W. Scheller, and C. Meyburg. 1995. Migration and Wintering of the Lesser Spotted Eagle (Aquila Pomarina)- A Study by Means of Satellite Telemetry. *Journal fur Ornithologie* 136: 401-422.

Midgley, G.F., L. Hannah, D. Millar, M.C. Rutherford, and L.W. Powrie. 2002. Assessing the vulnerability of species richness to anthropogenic climate change in a biodiversity hotspot. *Global Ecology and Biogeography* 11: 445-451.

Miller, M.W. and T.D. Nudds. 1996. Prairie landscape change and flooding in the Mississippi River Valley. *Conservation Biology* 10: 846-853.

Milsom, T.P., S.D. Langton, W.K. Parkin, S. Peel, J.D. Bishop, J.D. Hart, and N.P. Moore. 2000. Habitat models of bird species' distribution: an aid to the management of coastal grazing marshes. *Journal of Applied Ecology* 37: 706-727.

Minchin, P.R. 1987. An evaluation of the relative robustness of techniques for ecological ordination. *Vegetatio* 69: 89-107.

Moore, Frank R., and D.A. Aborn. 2000. Mechanisms of en route habitat selection : How do migrants make habitat decisions during stopover? *Studies in Avian Biology* 20: 34-42.

Naugle, D., R. Johnson, M. Estey, and K. Higgins. 2001. A landscape approach to conserving wetland bird habitat in the prairie pothole region of eastern South Dakota. *Wetlands* 21: 1-17.

Neave, H.M., R.B. Cunningham, T.W. Norton, and H.A. Nix. 1996. Biological inventory for conservation evaluation III. Relationships between birds, vegetation and environmental attributes in southern Australia. *Forest Ecology and Management* 85: 197-218.

Newman, S.H., S.A. Iverson, J.Y. Takekawa, M. Gilbert, D.J. Prosser, N. Batbayar, T. Natsagdorj, and D.C. Douglas. 2009. Migration of Whooper Swans and Outbreaks of Highly Pathogenic Avian Influenza H5N1 Virus in Eastern Asia. *Plos One*.

Oksanen, J., F.G. Blanchet, R. Kindt, P. Legendre, R.B. O'Hara, G.L. Simpson, P. Solymos, M. Henry, H. Stevens, H.Wagner, G.A. 2011. vegan: Community Ecology r Package

Pantin-Jackwood, M. J., D,E, Swayne. 2009. Pathogenesis and pathobiology of avian influenza virus infection in birds. Revue Scientifique et Technique-Office International des Epizooties 28: 113-136.

Paquette, J.H. Devries, R.B. Emery, D.W. Howerter, B.L. Joynt, T.P. Sankowski. 1997. Effects of Transmitters on Reproduction and Survival of Wild Mallards. *The Journal of Wildlife Management* 61: 953-961.

Parham, P.E., and E. Michael. 2010. "Modeling the Effects of Weather and Climate Change on Malaria Transmission. *Environmental health perspectives* 118 (5) (May): 620-626.

Park, P. 2004. Assessment and management of invasive alien predators. Ecology and Society 9: 12.

Parmesan, C. 2006. Ecological and Evolutionary Responses to Recent Climate Change. *Annual Review of Ecology, Evolution, and Systematics*: 637-669.

Parmesan, C., and G. Yohe. 2003. A globally coherent fingerprint of climate change impacts across natural systems. *Nature* 421: 37-42.

Pearson, R.G., and T.P. Dawson. 2003. Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Global Ecology & Biogeography* 12: 361-371.

Pearson, R.G., C.J. Raxworthy, M. Nakamura, and A.T. Peterson. 2007. Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *Journal of Biogeography* 34: 102-117.

Peters, R.L., and T.E. Lovejoy. 1994. *Global Warming and Biological Diversity*. Yale University Press, February 23.

Phillips, S.J., and M. Dudík. 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography* 31: 161-175.

Phillips, S.J., R.P. Anderson, and R.E. Schapire. 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling* 190: 231-259.

Phillips, S.J., M. Dudík, and R.E. Schapire. 2004. A maximum entropy approach to species distribution modeling. *Proceedings of the twenty-first international conference on Machine learning*, 83. Banff, Alberta, Canada: ACM.

Post, D.M., J.P. Taylor, J.F. Kitchell, M.H. Olson, D.E. Schindler, and B.R. Herwig. 1998. The role of migratory waterfowl as nutrient vectors in a managed wetland. *Conservation Biology* 12: 910-920.

Preuss, N.O. 2001. Hans Christian Cornelius Mortensen: Aspects of his life and of the history of bird ringing. *Aredea-Wageningen* 89: 1-6.

PRISM Climate Group. 2004. "PRISM Climate Group." *Oregon State University*. http://prism.oregonstate.edu.

R Development Core Team. 2011. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Randin, C.F, T. Dirnböck, S. Dullinger, N.E. Zimmermann, M. Zappa, and A. Guisan. 2006. Are niche-based species distribution models transferable in space? *Journal of Biogeography* 33: 1689-1703.

Ronka M.T.H., C.L.V. Saari, E.A. Lehikoinen, J. Suomela, and K. Hakkila. 2005. Environmental changes and population trends of breeding waterfowl in northern Baltic Sea. *Annales Zoologici Fennici* 42: 587-602.

Rubidge, E.M., W.B. Monahan, J.L. Parra, S.E. Camerson, and J.S. Brashares. 2011. The role of climate, habitat, and species co-occurrence as drivers of change in small mammal distributions over the past century. *Global Change Biology* 17: 696-708.

Schaub, M., R. Pradel, L. Jenni, and J. Lebreton. 2001. Migrating Birds Stop Over Longer than Usually Thought: An Improved Capture-Recapture Analysis. *Ecology* 82: 852-859.

Schneider, S.H. 1993. Biotic interactions and global change. In *Biotic interactions and global change.* Sunderland, Massachusetts: Sinauer.

Skagen, S.K., D.A. Granfors, and C.P. Melcher. 2008. On determining the signigicanve of ephemeral continental wetlands to North American migratory shorebirds. *The Auk* 125: 20-29.

Sorenson, L.G., R. Goldberg, T.L. Root, and M.G. Anderson. 1998. Potential Effects of Global Warming on Waterfowl Populations Breeding in the Northern Great Plains. *Climatatic Change* 40: 343-369.

Stevens, D.L., and S.F. Jensen. 2007. Sample Design, Execution, and Analysis for Wetland Assessment. *Wetlands* 27: 515-523.

Swets, J.A. 1988. Measuring the accuracy of diagnostic systems. *Science* 240: 1285-1293.

Takekawa, J.Y., S.H. Newman, X.M. Xiao, D.J. Prosser, K.A. Spragens, E.C. Palm, B.P. Yan, et al. 2010. Migration of Waterfowl in the East Asian Flyway and Spatial Relationship to HPAI H5N1 Outbreaks. *Avian Diseases* 54: 466-476.

Tamisier, A., A. Bechet, G. Jarry, J.C. Lefeuvre, and Y. Le Maho. 2003. Effects of hunting disturbance on waterbirds. A review of literature. *Revue D Ecologie-La Terre et la Vie*: 435-449.

Tankersley, R., and K. Orvis. 2003. Modeling the Geography of Migratory Pathways and Stopover Habitats for Neotropical Migratory Birds. *Conservation Ecology* 7: 7.

The Cornell Lab of Ornithology. 2011. "All About Birds." www.allaboutbirds.org.

Thomas, C.D., E.J. Bodsworth, R.J. Wilson, A.D. Simmons, Z.G. Davies, M. Musche, and L. Conradt. 2001. Ecological and evolutionary processes at expanding range margins. *Nature* 411: 577-581.

Thuiller, W., S. Lavorel, and M.B. Araújo. 2005. Niche properties and geographical extent as predictors of species sensitivity to climate change. *Global Ecology & Biogeography* 14: 347-357.

Tiner, R.W. 1984. Geographically Isolated Wetlands of the United States. *Wetlands* 23: 494-516.

Travis, J.M.J. 2003. Climate change and habitat destruction: a deadly anthropogenic cocktail. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 270: 467 -473.

VanDerWal, J., L.P. Shoo, C. Graham, and S. E. Williams. 2009. Selecting pseudo-absence data for presence-only distribution modeling: How far should you stray from what you know? *Ecological Modelling* 220: 589-594.

Velde, H. te. 1980. A Few Remarks upon the Religious Significance of Animals in Ancient Egypt. *Numen* 27 (1) (June 1): 76-82.

Visser, M.E. 2008. Keeping up with a warming world; assessing the rate of adaptation to climate change. *Proceedings of the Royal Society B: Biological Sciences* 275: 649 -659.

Walters, M. P. 2003. A Concise History of Ornithology. Book. http://nhm.openrepository.com/nhm/handle/10141/64051.

Ward, M.P., D.N. Maftei, C.L. Apostu, and A.R. Suru. 2009. Association Between Outbreaks of Highly Pathogenic Avian Influenza Subtype H5N1 and Migratory Waterfowl (Family Anatidae) Populations. *Zoonoses & Public Health* 56: 1-9.

Weber, T.P., A.I. Alasdair, and B.J. Ens. 1999. Consequences of habitat loss at migratory stopover sites: a theoretical investigation. *Journal of Avian Biology* 30: 416-426.

Wells, K., M. Suedkamp, B.E. Washburn, J.J. Millspaugh, M.R. Ryan, and M.W. Hubbard. 2003. Effects of Radio-Transmitters on Fecal Glucocorticoid Levels in Captive Dickcissels. *The Condor* 105: 805-810.

Williams, B.K., M.D. Koneff, and D.A. Smith. 1999. Evaluation of waterfowl conservation under the North American Waterfowl Management Plan. *Journal of Wildlife Management* 63: 417-440.

Wisz, M.S., R.J. Hijmans, J. Li, A.T. Peterson, C.H. Graham, A. Guisan, and NCEAS Predicting Species Distributions Working Group†. 2008. Effects of sample size on the performance of species distribution models. *Diversity and Distributions* 14: 763-773.

Yong, W., D.M. Finch, F.R. Moore, and J.F. Kelly. 1998. Stopover Ecology and Habitat Use of Migratory Wilson's Warblers. *The Auk* 115: 829-842.

Vita

Betty Jane Kreakie attended Ravenna High School in Ravenna, Ohio. In 1997, she began Kent State University in Kent, Ohio. Betty transferred to The Ohio State University in Columbus, Ohio, and graduated in 2002 with a Bachelor of Science in natural resources and urban forestry. During her time at The Ohio State University, she interned at the Franklin Soil and Water Conservation District and Dawes Arboretum. In 2002, Betty moved to North Carolina and worked for the Catena Group. Her time at the Catena Group was spent predominately working on a comprehensive biological survey of the Great Smoky Mountains National Park. In 2004, she began working on a Masters of Environmental Management at Duke's Nicholas School of the Environment. Her research focused on examining the vegetation community composition of South Florida's tropical hardwood hammocks. During this time, she was also employed at the Key West Tropical Forest and Botanical Garden to develop an environmental education curriculum for Monroe County, Florida. After graduating Duke in 2006, she worked for the North Carolina Natural Heritage Program mapping rare floristic species of North Carolina. In August 2006, she entered the University of Texas at Austin to begin a Doctor of Philosophy in Ecology, Evolution, and Behavior. Upon completion of her doctorate, Betty will be employed as a spatial population ecologist at the United States Environmental Protection Agency.

Permanent Address: b.kreakie@gmail.com

This manuscript was typed by the author.